

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2005, 23:17:42 ; Search time 101 Seconds
(without alignments)
3127.531 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVVEVAGSWMAQEREDILM.....TSGPCLGGLHLESSQPPPGF 549

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	549	1 TBC3_HUMAN	Q81zpl1 homo sapien
2	2971	99.6	549	2 Q6IPX1	Q6IPX11 homo sapien
3	2971	99.6	549	2 AAH71680	AAH716801 homo sapi
4	2945	98.7	549	2 Q6DHY5	Q6DHY51 homo sapien
5	2188.5	73.4	1406	1 UBP6_HUMAN	P351251 homo sapien
6	1544.5	51.8	376	2 Q15635	Q156351 homo sapien
7	1479	49.6	346	2 Q6PD72	Q6PD721 homo sapien
8	1479	49.6	346	2 AAH58890	AAH588901 homo sapi
9	1477.5	49.5	376	2 Q6PIJ2	Q6PIJ21 homo sapien
10	1477.5	49.5	376	2 AAH33670	AAH336701 homo sapi
11	1346	45.1	291	2 Q6DCB4	Q6DCB41 homo sapien
12	655	22.0	819	1 US6L_MOUSE	Q80xc31 mus. musculu
13	655	22.0	841	2 BAC97847	BAC978471 mus muscu
14	651	21.8	828	1 US6L_HUMAN	Q927381 homo sapien
15	651	21.8	838	2 BAA02807	BAA028071 homo sapi
16	603.5	20.2	250	2 Q86UD7	Q86UD71 homo sapien
17	602	20.2	224	2 Q8NCR5	Q8NCR51 homo sapien
18	576	19.3	1085	2 Q9U2D8	Q9U2D81 caenorhabdi
19	576	19.3	1085	2 CAB60374	CAB603741 caenorhab
20	574.5	19.3	1021	2 Q7K711	Q7K7111 caenorhabdi
21	574.5	19.3	1021	2 CAE45741	CAE457411 caenorhab
22	530	17.8	391	2 Q7QFV5	Q7QFV51 anopheles g
23	500.5	16.8	498	2 Q7PPV8	Q7PPV81 anopheles g
24	455	15.3	457	2 Q18357	Q183571 drosophila
25	455	15.3	485	2 Q95RE0	Q95RE01 drosophila
26	455	15.3	571	2 Q7KR75	Q7KR751 drosophila
27	323.5	10.8	356	2 Q81347	Q813471 arabidopsis
28	321.5	10.8	537	2 Q8BHL3	Q8BHL31 m mus muscu
29	321.5	10.8	537	2 Q91XR3	Q91XR31 mus musculu
30	321.5	10.8	655	2 Q6GQW9	Q6GQW91 mus musculu
31	320.5	10.7	533	2 Q9H8Z2	Q9H8Z21 homo sapien

32	320.5	10.7	584	2 Q6IN54	Q6IN541 homo sapien
33	320.5	10.7	584	2 AAH72453	AAH724531 homo sapi
34	320.5	10.7	622	2 Q6P530	Q6P5301 homo sapien
35	320.5	10.7	622	2 AAH63112	AAH631121 homo sapi
36	319	10.7	353	2 Q8LAF3	Q8LAF31 arabidopsis
37	317.5	10.6	304	2 Q9M894	Q9M8941 arabidopsis
38	306.5	10.3	532	2 Q9Y1V8	Q9Y1V81 halocynthia
39	305.5	10.2	500	1 TB10_MOUSE	P588021 mus musculu
40	302	10.1	298	2 Q8L756	Q8L7561 mus musculu
41	301	10.1	446	2 Q8IV04	Q8IV041 homo sapien
42	301	10.1	446	2 AAH62999	AAH629991 homo sapi
43	301	10.1	450	2 AAH36873	AAH368731 homo sapi
44	300	10.1	500	2 Q8CI62	Q8CI621 mus musculu
45	299.5	10.0	508	1 TB10_HUMAN	Q9bxi61 homo sapien

ALIGNMENTS

RESULT 1
TBC3_HUMAN
ID TBC3_HUMAN STANDARD; PRT; 549 AA.
AC Q8IZP1; Q9H0B9; Q9UDD4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE TBC1 domain family member 3 (Rab GTPase-activating protein PRC17)
DE (Prostate cancer gene 17 protein) (TRE17 alpha protein).
GN Name=TBC1D3; Synonyms=PRC17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND DISEASE.
RX MEDLINE=22246263; PubMed=12359748;
RA Pei L., Peng Y., Yang Y., Ling X.B., Van Eyrndhoven W.G., Nguyen K.C.,
RA Rubin M., Hoey T., Powers S., Li J.;
RT "PRC17, a novel oncogene encoding a Rab GTPase-activating protein, is
RT amplified in prostate cancer."
RL Cancer Res. 62:5420-5424(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasl S.,
RA Ansgorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Pousetka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RL Genome Res. 11:422-435(2001).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RX MEDLINE=93228825; PubMed=8471161;
RA Onno M., Nakamura T., Mariage-Samson R., Hillova J., Hill M.;
RT "Human TRE17 oncogene is generated from a family of homologous
RT polymorphic sequences by single-base changes."
RL DNA Cell Biol. 12:107-118(1993).
RN [4]
RP SEQUENCE OF 1-52 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94010310; PubMed=8406013;
RA Onno M., Nakamura T., Hillova J., Hill M.;
RT "Identification of novel sequences in the repertoire of hypervariable
RT TRE17 genes from immortalized normal and malignant human
RT keratinocytes."
RL Gene 131:209-215(1993).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=22506395; PubMed=12604796; DOI=10.1073/pnas.0437015100;
RA Paulding C.A., Ruvoilo M., Haber D.A.;

RT "The Tre2 (USP6) oncogene is a hominoid-specific gene."
RL Proc. Natl. Acad. Sci. U.S.A. 100:2507-2511(2003).
CC -1- FUNCTION: Acts as a GTPase activating protein for RAB5. Does not
CC act on RAB4 or RAB11.
CC -1- TISSUE SPECIFICITY: Expressed in liver, skeletal muscle, kidney,
CC pancreas, spleen, testis, ovary, small intestine and peripheral
CC blood leukocytes. Overexpressed in prostate cancers.
CC -1- DISEASE: May be involved in forms of prostate cancers.
CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF540953; AAN3117.1; -.
DR EMBL; AL136860; CAB66794.1; -.
DR EMBL; X71377; CAB94197.1; -.
DR Genew; HGNC:19031; TBC1D3.
DR MIM; 607741; -.
DR InterPro; IPR00195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
DR GTPase activation; Proto-oncogene.
DR DOMAIN 101 293 Rab-GAP TBC.
FT CONFLICT 117 117 I -> T (in Ref. 2).
FT CONFLICT 137 137 R -> K (in Ref. 2).
FT CONFLICT 354 354 Q -> K (in Ref. 2).
FT CONFLICT 526 526 P -> Q (in Ref. 2).
SQ SEQUENCE 549 AA; 62168 MW; 831DA747D0B6F400 CRC64;

Query Match
Best Local Similarity 100.0%; Score 2983; DB 1; Length 549;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGI VHETELPPL 60
DB 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGI VHETELPPL 60
QY 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLNPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIIFRDRYGTQKORELHILLAYEEX 180
DB 121 KLNPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIIFRDRYGTQKORELHILLAYEEX 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOQE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOQE 240
QY 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQOKRLITKTSRCGPMWAFPCNRFVDTWARDEDTVLKLRA SMKKLTRKQGDLPFP 360
DB 301 TRIAFKVQOKRLITKTSRCGPMWAFPCNRFVDTWARDEDTVLKLRA SMKKLTRKQGDLPFP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGPGSGWRFLQWNSMPRLPTDLVEGPMFRHYDPRQSCWV 480
DB 421 VREDTYPVGTQGVSPALAOGPGSGWRFLQWNSMPRLPTDLVEGPMFRHYDPRQSCWV 480
QY 481 RAISQEQOLAPCWQAEHPAERVRSAFAAPSTDSQGTFFRARDEQPCAPTSGPCLCGLHL 540
DB 481 RAISQEQOLAPCWQAEHPAERVRSAFAAPSTDSQGTFFRARDEQPCAPTSGPCLCGLHL 540

DB 481 RAISQEQOLAPCWQAEHPAERVRSAFAAPSTDSQGTFFRARDEQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 2
Q6IPX1
ID Q6IPX1 PRELIMINARY; PRT; 549 AA.
AC Q6IPX1;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE TBC1 domain family, member 3.
GN Name=TBC1D3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maira P.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.,
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071680; AAH71680.1; -.
DR InterPro; IPR00195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
SQ SEQUENCE 549 AA; 62231 MW; 5657E5602B3655F5 CRC64;

Query Match
Best Local Similarity 99.6%; Score 2971; DB 2; Length 549;
Matches 546; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGI VHETELPPL 60
DB 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGI VHETELPPL 60
QY 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLNPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIIFRDRYGTQKORELHILLAYEEX 180
DB 121 KLNPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIIFRDRYGTQKORELHILLAYEEX 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOQE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOQE 240

Db 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGPHSPNGSTVQGLQDQE 240
QY 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQOKRLTTSRCGFWARFCNRFVDTWARDEDVTLKHLRASMKKLTRKQGDLP 360
Db 301 TRIAFKVQOKRLTTSRCGFWARFCNRFVDTWARDEDVTLKHLRASMKKLTRKQGDLP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
QY 481 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSOGTFFRARDEQCAPTSGPCLGHL 540
Db 481 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSOGTFFRARDEQCAPTSGPCLGHL 540
QY 541 ESSQFPPGF 549
Db 541 ESSQFPPGF 549

RESULT 3
AAH71680 PRELIMINARY; PRT; 549 AA.
ID AAH71680
AC AAH71680;
DT 01-JUN-2004 (TREMBlrel. 27, Created)
DT 01-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE TBC1 domain family, member 3.
GN TBC1D3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strauberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071680; AAH71680.1; -
SQ SEQUENCE 549 AA; 62231 MW; 5657E5602B3655F5 CRC64;

Query Match 99.6%; Score 2971; DB 2; Length 549;
Best Local Similarity 99.5%; Pred. No. 8.2e-179;
Matches 546; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVEVAGSWWAQEREDIMKEKGHRAGLPEDKPKPFRSYNNVNDHGIHETELPPL 60
Db 1 MDVEVAGSWWAQEREDIMKEKGHRAGLPEDKPKPFRSYNNVNDHGIHETELPPL 60
QY 61 TAREAKQIRREISRSKVMMDLGDEWEKYSSRKLIDRAYKGMNINRGPMVSVLNIEM 120
Db 61 TAREAKQIRREISRSKVMMDLGDEWEKYSSRKLIDRAYKGMNINRGPMVSVLNIEM 120
QY 121 KLKNPGRYQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDRYGTQRELLHLLAYEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDRYGTQRELLHLLAYEY 180
QY 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGSTVQGLQDQE 240
Db 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGSTVQGLQDQE 240
QY 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQOKRLTTSRCGFWARFCNRFVDTWARDEDVTLKHLRASMKKLTRKQGDLP 360
Db 301 TRIAFKVQOKRLTTSRCGFWARFCNRFVDTWARDEDVTLKHLRASMKKLTRKQGDLP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
QY 481 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSOGTFFRARDEQCAPTSGPCLGHL 540
Db 481 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSOGTFFRARDEQCAPTSGPCLGHL 540
QY 541 ESSQFPPGF 549
Db 541 ESSQFPPGF 549

RESULT 4
Q6DHYS PRELIMINARY; PRT; 549 AA.
ID Q6DHYS
AC Q6DHYS;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC075809, AAH75809.1; -.
KW Hypothetical protein.
SQ SEQUENCE 549 AA; 62304 MW; 401430A45C00D9C5 CRC64;

Query Match 98.7%; Score 2945; DB 2; Length 549;
Best Local Similarity 98.5%; Pred. No. 3.5e-177;
Matches 541; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVEVAGSWAQRREDIMKYEKGHAGLPEDKGPFRSYNNVNDHLGLVHETELPPL 60
Db 1 MDVEVAGSWAQRREDIMKYEKGHAGLPEDKGPFRSYNNVNDHLGLVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKWVDMLGWEKYSSRKLIDRAYKGMNIRGPMWSVLNITEEM 120
Db 61 TAREAKQIRREISRSKSKWVDMLGWEKYSSRKLIDRAYKGMNIRGPMWSVLNITEEM 120
QY 121 KLKNPGRYQIMKEKGKRSSEHIQIRIDVSGTLRKHIFRDRYGTQKRELIHLLAYEY 180
Db 121 KLKNPGRYQIMKEKGKRSSEHIQIRIDVSGTLRKHIFRDRYGTQKRELIHLLAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGSTVQGLQDQOE 240
Db 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGSTVQGLQDQOE 240
QY 241 HVVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMP 300
Db 241 HVVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMP 300
QY 301 TRIAFKYQOKRLTTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTRKQGDLP 360
Db 301 TRIAFKYQOKRLTTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTRKQGDLP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPGPPARFPPIWASPPRARSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPGPPARFPPIWASPPRARSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLDVEGPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLDVEGPMFRHYDFRQSCW 480
QY 481 RAISOEDQLAPCWAHHPAERVRSAPAPASTDSQGTFFARDEQPCAPTSGPCLCGLHL 540
Db 481 RAISOEDQLAPCWAHHPAERVRSAPAPASTDSQGTFFARDEQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPFPGF 549
Db 541 ESSQFPFPGF 549

RESULT 5
UBP6_HUMAN
ID UBP6_HUMAN STANDARD; PRT; 1406 AA.
AC P35125; Q15634; Q86WP6; Q81WT4;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase 6 (EC 3.1.2.15) (Ubiquitin
thiolesterase 6) (Ubiquitin-specific processing protease 6)
DE (Deubiquitinating enzyme 6) (Proto-oncogene TRB-2).
GN Name=USP6; Synonyms=TRB2;

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.
RC TISSUE=Ewing sarcoma;
RX MEDLINE=92228503; PubMed=1565468;
RA Nakamura T., Hillova J., Mariage-Samson R., Onno M., Huebner K.,
RA Cannizzaro L.A., Boghosian-Sell L., Croce C.M., Hill M.;
RT "A novel transcriptional unit of the tre oncogene widely expressed in
RT human cancer cells.";
RL Oncogene 7:733-741(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING, TISSUE
RP SPECIFICITY, AND DISCUSSION OF TRE2 EVOLUTION.
RX MEDLINE=22506395; PubMed=12604796; DOI=10.1073/pnas.0437015100;
RA Paulding C.A., Ruvoilo M., Haber D.A.;
RT "The Tre2 (USP6) oncogene is a hominoid-specific gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:2507-2511(2003).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=94067315; PubMed=8247125;
RA Papa F.R., Hochstrasser M.;
RT "The yeast DOA4 gene encodes a deubiquitinating enzyme related to a
RT product of the human tre-2 oncogene.";
RL Nature 366:313-319(1993).
RN [4]
RP MUTAGENESIS OF THR-150 AND ARG-187.
RX PubMed=14521938;
RA Bizimungu C., De Neve N., Burny A., Bach S., Bontemps F.,
RA Portetelle D., Vandenbol M.;
RT "Expression in a RabGAP yeast mutant of two human homologues, one of
RT which is an oncogene.";
RL Biochem. Biophys. Res. Commun. 310:498-504(2003).
CC -1- FUNCTION: Has an ATP-independent isopeptidase activity, cleaving
CC at the C-terminus of the ubiquitin moiety. In vitro, isoform 2,
CC but not isoform 3, shows deubiquitinating activity.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P35125-1; Sequence=Displayed;
CC Name=2; Synonyms=213(ORF2);
CC IsoId=P35125-2; Sequence=VSP_010878, VSP_010879;
CC Name=3; Synonyms=210(ORF1), onctre210p;
CC IsoId=P35125-3; Sequence=VSP_010880, VSP_010881;
CC Note=Was shown to be tumorigenic in transfectected mice and seems
CC not to act as GTPase activating protein;
CC -1- TISSUE SPECIFICITY: Testis specific. Expressed in various cancer
CC cell lines.
CC -1- DISEASE: Involved in tumorigenesis.
CC -1- MISCELLANEOUS: The USP6 gene only exists in the primate lineage.
CC -1- SIMILARITY: Belongs to peptidase family C19.
CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi-sib.ch/announce/>
CC or send an email to license@ebi-sib.ch).
CC -----
DR EMBL; X63546; CAA45108.1; -.
DR EMBL; X63547; CAA45111.1; -.
DR EMBL; AY143550; AAN38838.1; -.
DR EMBL; AY163314; AAO21348.1; -.
DR PIR; S57867; S57867.
DR PIR; S57868; S22158.
DR PIR; S57874; S22155.
DR MEROPS; C19.009; -.

DR Genew; HGNC:12629; USP6.
DR MIM; 604334; -.
DR GO; GO:0005764; C:lysosome; TAS.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; TAS.
DR GO; GO:0003676; F:nucleic acid binding; TAS.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR Pfam; PF00443; UCH; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
KW Alternative splicing; Hydrolase; Multigene family; Proto-oncogene;
KW Thiol protease; Ubl conjugation pathway.
FT DOMAIN 100 292 Rab-GAP TBC.
FT ACT_SITE 755 755 By similarity.
FT ACT_SITE 1320 1320 By similarity.
FT ACT_SITE 1328 1328 By similarity.
FT VARSPLIC 1 317 Missing (in isoform 2).
FT VARSPLIC 318 359 /FTid=VSP_010878.
FT GLMARLRNQFDTWAMNDTVLKHLRASSTKKLTRKQGLPP
FT P -> MPQRLPHARQHTPLPLGSADYRRVSVRPOGPHRDP
FT KDSRDA (in isoform 2).
FT /FTid=VSP_010879.
FT NFPQDNQKVQLSV -> ISPLHLQMECSP (in
FT isoform 3).
FT /FTid=VSP_010880.
FT Missing (in isoform 3).
FT /FTid=VSP_010881.
FT T->R: Does not restore GAP activity in
FT yeast complementation assay.
FT R->Q: Does not restore GAP activity in
FT yeast complementation assay.
FT CONFLICT 475 475 W -> R (in Ref. 1).
FT CONFLICT 912 912 R -> Q (in Ref. 1).
FT CONFLICT 963 963 N -> I (in Ref. 1).
SQ SEQUENCE 1406 AA; 158657 MW; D3A6822CEB441DB3 CRC64;
Query Match 73.4%; Score 2188.5; DB 1; Length 1406;
Best Local Similarity 81.6%; Pred. No. 4.3e-129;
Matches 408; Conservative 32; Mismatches 59; Indels 1; Gaps 1;
QY 1 MDVVEVAGSWWAQEREDIIIMKYBKGRAGLPEDKGPKEPRSYNNNVHIGIVHETELPPL 60
DB 1 MDVVENADSLQAQERKDILMKYDKGHRAGLPEDKGPPEV-GINSSIDRFGLIHETELPPV 59
QY 61 TAREAKQIRREISRSKAVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIIEEM 120
DB 60 TAREAKKIRREMTRTSKWMEMLGEMETYSKSLIDRVYKGI PMNIRGPVMSVLNIQEI 119
QY 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQRELLHILAYEY 180
DB 120 KLKNPGRYQIMKEKGRSSEHIHIDLDVPTLTNRHVFPRDRYGAKQRELFYILAYSEY 179
QY 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDQOE 240
DB 180 NPEVGYCRDLISHITALFLLYLPEEDAFWALVOLLASERHSLPGFHS PNGGTVOGLQDQOE 239
QY 241 HVVATSQPKTMGHQDKDLQGCQSPGLCLIRLIDIGISLGLTLRLMDVYLVVEGEQALMPI 300
DB 240 HVPKPSQPKTMGHQDKEGLCGQCASLGLCLRLNLDIGISLGLTLRLMDVYLVVEGEQALMPI 299
QY 301 TRIAFKVOQKRLTKTSRCGFWARFCNRFVDTWARDEDTV LKHLRASMKKLTRKQGLPPP 360
DB 300 TSIALKVOQKRLTKTSRCGLMARLRNQFFDTWAMNDTV LKHLRASSTKKLTRKQGLPPP 359
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPGPPARFPRIW SASPPRARSSTPCPGA 420
DB 360 AKREGSLAPRPVPASRGKTLCKGYRQAPGPPAQFORP ICSASPPMASRSTPCPGA 419
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTD LDVEGPFRRHYDFRQSCWV 480

DB 420 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTD LDIGGPFRRHYDFEWSCWV 479
QY 481 RAISQEDQLAPCWAHPAE 500
DB 480 RAISQEDQLATCWAHPHCE 499
RESULT 6
Q15635
ID Q15635 PRELIMINARY; PRT; 376 AA.
AC Q15635;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Oncogene.
GN Name=trec;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Esivings' sarcoma;
RX MEDLINE=92228503; PubMed=1565468;
RA Nakamura T., Hillova J., Mariage-Samson R., Onno M., Huebner K.,
RA Cannizzaro L.A., Boghosian-Sell L., Croce C.M., Hill M.;
RT "A novel transcriptional unit of the trec oncogene widely expressed in
RT human cancer cells."
RL Oncogene 7:733-741(1992).
DR EMBL; X63547; CAA45110.1; -.
DR PIR; S57867; S57867.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
SQ SEQUENCE 376 AA; 43160 MW; F43A86D09F344117 CRC64;
Query Match 51.8%; Score 1544.5; DB 2; Length 376;
Best Local Similarity 77.9%; Pred. No. 3.1e-89;
Matches 293; Conservative 29; Mismatches 51; Indels 3; Gaps 2;
QY 1 MDVVEVAGSWWAQEREDIIIMKYBKGRAGLPEDKGPKEPRSYNNNVHIGIVHETELPPL 60
DB 1 MDVVENADSLQAQERKDILMKYDKGHRAGLPEDKGPPEV-GINSSIDRFGLIHETELPPV 59
QY 61 TAREAKQIRREISRSKAVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIIEEM 120
DB 60 TAREAKKIRREMTRTSKWMEMLGEMETYSKSLIDRVYKGI PMNIRGPVMSVLNIQEI 119
QY 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQRELLHILAYEY 180
DB 120 KLKNPGRYQIMKEKGRSSEHIHIDLDVPTLTNRHVFPRDRYGAKQRELFYILAYSEY 179
QY 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDQOE 240
DB 180 NPEVGYCRDLISHITALFLLYLPEEDAFWALVOLLASERHSLPGFHS PNGGTVOGLQDQOE 239
QY 241 HVVATSQPKTMGHQDKDLQGCQSPGLCLIRLIDIGISLGLTLRLMDVYLVVEGEQALMPI 300
DB 240 HVPKPSQPKTMGHQDKEGLCGQCASLGLCLRLNLDIGISLGLTLRLMDVYLVVEGEQALMPI 299
QY 301 TRIAFKVOQKRLTKTSRCGFWARFCNRFVDTWARDEDTV LKHLRASMKKLTRKQGLPPP 360
DB 300 TSIALKVOQKRLTKTSRCGLMARLRNQFFDTWAMNDTV LKHLRASSTKKLTRKQGLPPP 359
QY 361 AKPEQSS--ASRPVP 374
DB 360 GPTALGRRCVAGSPQP 375
RESULT 7
Q6PD72

```
ID Q6PD72 PRELIMINARY; PRT; 346 AA.
AC Q6PD72;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedln T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058890; AAH58890.1; -.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
KW Hypothetical protein.
SQ SEQUENCE 346 AA; 39686 MW; AD8F92563358FCBF CRC64;

Query Match
Best Local Similarity 49.6%; Score 1479; DB 2; Length 346;
Matches 276; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAOEREDIIIMKYEGHRAGLPEDKGPFRSYNNNVNVDHLGIHETELPPL 60
DB 1 MDVVEVAGSWMWAOEREDIIIMKYEGHRAGLPEDKGPFRSYNNNVNVDHLGIHETELPPL 60
QY 61 TAREVKQIRREISRSKSKWVMDMLGDWEKYKSSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREVKQIRREISRSKSKWVMDMLGDWEKYKSSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
DB 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
QY 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
DB 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
QY 241 HVVATSLPKTMWHQDKDLCGQCSPLGCLIRILIDGISTGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSLPKTMWHQDKDLCGQCSPLGCLIRILIDGISTGLTLRLMDVYLVEGEQALMPI 300
QY 301 TR 302
DB 301 TR 302
```

```
DB 301 TK 302

RESULT 8
ID AAH58890 PRELIMINARY; PRT; 346 AA.
AC AAH58890;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedln T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058890; AAH58890.1; -.
KW Hypothetical protein.
SQ SEQUENCE 346 AA; 39686 MW; AD8F92563358FCBF CRC64;

Query Match
Best Local Similarity 49.6%; Score 1479; DB 2; Length 346;
Matches 276; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAOEREDIIIMKYEGHRAGLPEDKGPFRSYNNNVNVDHLGIHETELPPL 60
DB 1 MDVVEVAGSWMWAOEREDIIIMKYEGHRAGLPEDKGPFRSYNNNVNVDHLGIHETELPPL 60
QY 61 TAREVKQIRREISRSKSKWVMDMLGDWEKYKSSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREVKQIRREISRSKSKWVMDMLGDWEKYKSSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
DB 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
QY 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
DB 121 KLKNPGRYQIMKEKGKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLASERHSLQGFHSPNGGTVOGLDQOE 240
QY 241 HVVATSLPKTMWHQDKDLCGQCSPLGCLIRILIDGISTGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSLPKTMWHQDKDLCGQCSPLGCLIRILIDGISTGLTLRLMDVYLVEGEQALMPI 300
QY 301 TR 302
DB 301 TR 302
```

Db 301 TK 302

RESULT 9

Q6PIJ2 PRELIMINARY; PRT; 376 AA.
ID Q6PIJ2
AC Q6PIJ2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, last annotation update)
DE Similar to TBC1 domain family, member 3.
GN Name=MGc44903;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033670; AAH33670.1; -
DR InterPro; IPR00195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PSS0086; TBC RABGAP; 1.
SQ SEQUENCE 376 AA; 42391 MW; F3A9AC94995B9900 CRC64;

Query Match 49.5%; Score 1477.5; DB 2; Length 376;
Best Local Similarity 72.7%; Pred. No. 5.1e-85;
Matches 301; Conservative 12; Mismatches 40; Indels 61; Gaps 8;

QY 1 MDVEVAGSWWAQEREDIIIMYKKGHRAGLPEDKGPFRSYNNVNDHLGIIVHETELPPL 60
DB 1 MDVEVAGSWWAQEREDIIIMYKKGHRAGLPEDKGPFRSYNNVNDHLGIIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSKSKRLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
DB 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSKSKRLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
QY 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKRELLHILLAYEEX 180
DB 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKRELLHILLAYEEX 180
QY 161 NPEVGYCRDLSHIALFLLYLPBEDAFWALVOLLASERHSLQGFHSPNGCTVQGLQDQOE 240
DB 161 NPEVGYCRDLSHIALFLLYLPBEDAFWALVOLLASERHSLQGFHSPNGCTVQGLQDQOE 240

QY 241 HVVATSQPTMGHQDKDLCGQCSPLGLIRILIDIGISLGLTLRLMDVYLVEGE---QA 296
DB 241 HVVATSQSKTMGHQDKDLCGQCSPLGLIRILIDGV-----RRHRETLAQGPSLPCSA 294

QY 297 IMP-----ITRIAPKVQOKRLTKTSRCGPWFARFCNRFVDTWARDED 337
DB 295 LIPQPGLAHSQPTGSGSGSPKDTQAKPSAQE-----GSSQGNWGSGBPASWAD-WAR--- 346
QY 338 TVLKHLRASMKKLTKRQGDLPFPAPKEQSSASRPVPASRGKTLCKDROQAP 391
DB 347 -----TRLERAQGSLLKPWASP-----SLQEPHPHS-----NECP 376

RESULT 10

AAH33670 PRELIMINARY; PRT; 376 AA.
ID AAH33670
AC AAH33670;
DT 01-JUN-2004 (TReMBLrel. 27, Created)
DT 01-JUN-2004 (TReMBLrel. 27, last sequence update)
DT 01-JUN-2004 (TReMBLrel. 27, last annotation update)
DE TBC1D3 protein.
GN TBC1D3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033670; AAH33670.1; -
SQ SEQUENCE 376 AA; 42391 MW; F3A9AC94995B9900 CRC64;

Query Match 49.5%; Score 1477.5; DB 2; Length 376;
Best Local Similarity 72.7%; Pred. No. 5.1e-85;
Matches 301; Conservative 12; Mismatches 40; Indels 61; Gaps 8;

QY 1 MDVEVAGSWWAQEREDIIIMYKKGHRAGLPEDKGPFRSYNNVNDHLGIIVHETELPPL 60
DB 1 MDVEVAGSWWAQEREDIIIMYKKGHRAGLPEDKGPFRSYNNVNDHLGIIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSKSKRLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
DB 61 TAREAKQIRREISRSKSKVNDMLGDWEKYSKSKRLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
QY 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKRELLHILLAYEEX 180
DB 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKRELLHILLAYEEX 180

QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOOE 240
QY 241 HVVATSOPTKMGHODKDCGQCSPLGCLIRILLIDGISTGLTLRLMDVYLVEGE---QA 296
DB 241 HVVATSOPTKMGHODKDCGQCSPLGCLIRILLIDGISTGLTLRLMDVYLVEGE---QA 296
QY 297 LMP-----ITRIAFKVQOKRLTTSRCGPWFARFENRFDVTWARDDED 337
DB 295 LLPQPGGLAHSQPTGSGSGSPKDTQAKPSAQE-----GSSQGNMGSGSPASWAD-WAR---- 346
QY 338 TVLKLRLASMKLTKRQGDLPFPAKPEQSSASRPVPASRGKTLCKGDRQAPP 391
DB 347 -----TRLERAQSGSLKWPASP-----SLQEPHPHS-----NECP 376

RESULT 11

Q6DCB4
ID Q6DCB4 PRELIMINARY; PRT; 291 AA.
AC Q6DCB4;
DT 01-OCT-2004 (Tremblrel. 28, Created)
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC078140; AAH78140.1; -
KW Hypothetical protein.
SQ SEQUENCE 291 AA; 33400 MW; D21C628DD436BCBA CRC64;

Query Match 45.1%; Score 1346; DB 2; Length 291;
Best Local Similarity 99.2%; Pred. No. 7.2e-77;
Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWAQEREDIIIMKYKGHRAGLPEDKGPKPFRRSYNNVNDHGLYHETELPPL 60
DB 1 MDVVEVAGSWAQEREDIIIMKYKGHRAGLPEDKGPKPFRRSYNNVNDHGLYHETELPPL 60
QY 61 TAREAKQIRREISRSKSWVMDLGDWEKYSRKLIIDRAYKGMPPNIRGPMWSVLNTEEM 120
DB 61 TAREAKQIRREISRSKSWVMDLGDWEKYSRKLIIDRAYKGMPPNIRGPMWSVLNTEEM 120

DB 61 TAREAKQIRREISRSKSWVMDLGDWEKYSRKLIIDRAYKGMPPNIRGPMWSVLNTEEM 120
QY 121 KLNKPGRYQIMKEKGSSSEHIQIRIDPDSGTLRKHIFFRDRYGTQRELLHILAYEY 180
DB 121 KMNKPGRYQIMKEKGSSSEHIQIRIDPDSGTLRKHIFFRDRYGTQRELLHILAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOOE 240
DB 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOOE 240
QY 241 HVVATSOPTKMGHQ 254
DB 241 HVVATSOPTKMGHQ 254

RESULT 12

US6L_MOUSE
ID US6L_MOUSE STANDARD; PRT; 819 AA.
AC Q80XC3;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DE USP6 N-terminal like protein.
GN Name=Usp6rl; Synonyms=Kiaa0019;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryonic tail;
RX PubMed=14621295;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT I1. The complete nucleotide sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [1]
RP FUNCTION: Acts as a GTPase activating protein for RAB3A. Involved
RP in receptor trafficking. In complex with EPS8 inhibits
RP internalization of EGFR (By similarity).
CC SUBUNIT: Interacts with EPS8 (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q80XC3-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q80XC3-2; Sequence=VSP_011153;

[illegible][illegible]

RA Matoskova B.;
RT "RN-tre identifies a family of tre-related proteins displaying a novel
RT potential protein binding domain.",
RL Oncogene 12:2563-2571(1996).
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
RT cell line KG-1.";
RL DNA Res. 1:27-35(1994).
RN [3]

RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257, PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman A., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]

RP FUNCTION, AND MUTAGENESIS OF ARG-106; ASP-147 AND ARG-150.
RX PubMed=11099046; DOI=10.1038/35042605;
RA Lanzetti L., Rybin V., Malabarba M.G., Christoforidis S., Scita G.,
RA Zerial M., Di Fiore P.P.;
RT "The Eph8 protein coordinates EGF receptor signalling through Rac and
RT trafficking through Rabs.";
RL Nature 408:374-377(2000).
CC -1- FUNCTION: Acts as a GTPase activating protein for RAB5A. Involved
CC in receptor trafficking. In complex with EPS8 inhibits
CC internalization of EGFR.
CC -1- SUBUNIT: Interacts with EPS8.
CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; D13644; BA002807.2; ALT_INIT.
CC EMBL; BC042943; AAH42943.1; -
CC Genew; HGNC:16858; USP6NL.
CC MIM; 605405; -
CC InterPro; IPR000195; RabGAP_TBC.
CC Pfam; PF00566; TBC; 1.
CC SMART; SM00164; TBC; 1.
CC PROSITE; PSS0086; TBC_RABGAP; 1.
CC KW GTPase activation.
CC FT DOMAIN 100 292 Rab-GAP TBC.
CC FT 106 106 R->A: Loss of GAP activity on RAB5A.

```

FT      MUTAGEN      147      147      D->A: Loss of GAP activity on RAB5A.
FT      MOTAGEN      150      150      R->A: Loss of GAP activity on RAB5A.
SQ      SEQUENCE      828 AA; 94104 MW; A5E3DBF0FEE04E78 CRC64;

Query Match
Best Local Similarity 21.8%; Score 651; DB 1; Length 828;
Matches 164; Conservative 72; Mismatches 172; Indels 80; Gaps 14

QY      12 AOEREDIIIMKEYEGHRAGLPEDKPKPRFSYNNDVHLGIYHETELPPLTAREAKOIRRE 71
Db       12 AQERAIEIVAKYDRGREGAEIPEWEDADYLVI-KVTDRFGFLHEELPDHNVAVEROKHLE 70
QY      72 ISRSKSVMDLGDMEKYKSSRLIDRAYKGEMNIRGPMSVLINIEEMKLKNRGYQIM 131
Db       71 IERTTKMLKWLKGWEKYKTEKEFHRIYKGIPLQLRGEWALLLEIPKMEETRDLYSKL 130
QY      132 KEKGKRSSSEHIQRIDRVSGTLRKHFIFRDRTGTQORELIHLLAYEEYNPEVGCRDLS 191
Db       131 KHRAGCGSPDIRQIDLDVNRTFRDHIMFRDRYGVKQSLPHVLAAYSINYNTVEVGCQWS 190
QY      192 HIAALFLLYLPEDDAFWALVOLLASERSHLQGFPHPNGTGVOG-----LODOEHYVAT 245
Db       191 QITALLIMYMNEEDAFWALVKLFSGPKHAMHGFF-----VGFEPKILRFQEHHEKITLNK 244
QY      246 SQPTMGHQDKDLCGQCSPLGLIRILIDISLGTLRLWDVYLVEGEQALMPITRIAF 305
Db       245 FLSKLKQHLDSEQEIYTSFYTMKMFPQCFLDRTPTLINLRIMDIYIFEGERULTAMSYTIL 304
QY      306 KVQOKRLTTKSRCGPMARFCNRFDVTWARD---EDTVLKHILRASMKULTRKQGDLPPIPA 361
Db       305 KLHKHLMKLS---MEEIVEFFQETLAKDPFFEDDFVTEQLQISMTELKRAKIDLPEPG 360
QY      362 K-----PEQS-----SASRPVPASRGKTLCKGDRQAP----- 390
Db       361 KEDEYPRKKPLGQLPPELQSWGHHLSNGORSVGRPSPLASG-----RRESGAPRRRHHS 415
QY      391 PGPARFPRPIWSASPPRAPRSSSTPCPGAVREDITYPVGTQG-----VSPPALAQGGPQG 445
Db       416 PHPOSRTGTP-ERAQPPIRRK-----SVEESKSKLKDEADFORKLPS-----GPQD 459
QY      446 SWREFLOWN 453
Db       460 SSR-QYN 465

RESULT 15
BAA02807 PRELIMINARY; PRT; 838 AA.
AC BAA02807;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE KIAA0019 protein (Fragment).
GN KIAA0019.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RT "Prediction of the coding sequences unidentified human genes. I. The
RT coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RL cell line KG-1."
RN DNA Res. 1:27-35(1994).
RP [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051389; PubMed=7584028;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K., Tabata S.;
```

RT "Prediction of the coding sequences of unidentified human genes 1. The
RT coding sequences of 40 new genes(KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1 (supplement).";
RL DNA Res. 1:47-56(1994).

RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96293402; PubMed=8700515;
RA Wong W.T., Seki N., Nagase T., Nomura N., Robbins K.C., Di Fiore P.P.,
RA Matoskova B.;
RT "RN-tre identifies a family of tre-related proteins displaying a novel
RT potential protein binding domain.";
RL Oncogene 12:2563-2571(1996).

RN [4]
RP SEQUENCE FROM N.A.
RA Ohara O., Nagase T., Kikuno R., Nomura N.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D13644; BAA02807.2; -.
FT NON TER 1
SQ SEQUENCE 838 AA; 95182 MW; EF5E4E08D7099B07 CRC64;

Query Match 21.8%; Score 651; DB 2; Length 838;
Best Local Similarity 33.6%; Pred. No. 1.3e-32;
Matches 164; Conservative 72; Mismatches 172; Indels 80; Gaps 14;

QY 12 AQEREDIMKYEKGHRAGLPEDKGPFPSSYNNDHGIHETELPPLTAREAKQIRRE 71
DB 22 AQERAEIVAKYDRGREGAEIBPWEDADYLY-KVTDRFGLHEEELPDHNAVERQKHLE 80
QY 72 ISRSKQWVDMLGDMEKYKSSRKLIIDRAYGMPNIRGPMWSVLNIEEMKLNPGRYQIM 131
DB 81 IERTTKMLKMGWEKYKTEKFRRIYKGPLQLRGEVWALLLEIPKKEETRDLYSKL 140
QY 132 KEKGRSSSEHIQRIIDRVSGTLRKHIFPRDRYGTQRELBHLLAYEYNPEVGYCRDLS 191
DB 141 KHRAGCSPDIRQIDLDVNRITFRDHIMFRDRYGVKQSLFHVLAAYSINTEVGYCQGM 200
QY 192 HIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGTVQG-----LQDQOEHVAT 245
DB 201 QITALLMYNEDEDAFWALVKLFGSPKHAMGFF-----VQGPRLRFQEHHEKILNK 254
QY 246 SQPKTMGHODKDLGCGCSPLGLIRLIDGSLGLTLRLWDVYLVEGEQALMPITRIAP 305
DB 255 FLSKLKQHLDSQEIYTSFYTMKWFQCFDRTPFTLNLRIMDIYFEGERVLTAMSYTIL 314
QY 306 KVOQKRLTKTSRCGPWARFCNRFVDTWARD---EDTVLKHILRASMKL/TRKQGLPPPA 361
DB 315 KLHKHLMKLS---MELVVEFQETLAKDFFEDDFVIEQLQISMTLKRKAKLDLPEPG 370
QY 362 K-----PEQGS-----SASRPVPASRGKTLCKGDRQAP----- 390
DB 371 KEDEYPKKPLGQLPPELQSGVHHLNSGQSVGRPSPLASG-----RREGAPHRHENS 425
QY 391 PGPPARFPPIWSASPPRAPRSSTPCPGAVREDTYPVGTQG-----VSPALAQGGPQG 445
DB 426 PHPQSRGTGP-ERAQPPRRK-----SVEESKCLKDEADFORKLPS-----GPD 469
QY 446 SWRFLQWN 453
DB 470 SSR--QYN 475

Search completed: February 4, 2005, 05:10:05
Job time : 107 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 05:11:13 ; Search time 6811 Seconds
(without alignments)
3811.782 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983
Sequence: 1 MDVVEVAGSWWAQEREDII.....TSGPCLGHLHSSQFPFPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US10071838/runat_03022005_071218_187/app_query.fasta_1.711
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071838@CGN_1_1_3852@runat_03022005_071218_187 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	6	AX775945 Sequence
2	2983	100.0	1964	9	AF540953 Homo sapi
3	2975	99.7	1964	6	AX775947 Sequence
4	2971	99.6	2090	9	BC071680 Homo sapi

5	2963	99.3	2072	6	AX086847 Sequence
6	2963	99.3	2072	9	AL136860 Homo sapi
7	2945	98.7	2166	9	BC075809 Homo sapi
8	2906.5	97.4	2304	6	AX775943 Sequence
9	2898.5	97.2	2304	6	AX775941 Sequence
10	2849	95.5	2575	9	BC033670 Homo sapi
11	2835	95.0	2647	6	AX775951 Sequence
12	2827	94.8	2647	6	AX775949 Sequence
13	2827	94.8	2647	6	AX833552 Sequence
14	2827	94.8	2647	9	AK095385 Homo sapi
15	2732	91.6	2111	9	BC058890 Homo sapi
16	2626.5	88.0	1924	9	BC078140 Homo sapi
17	2595	87.0	3389	9	AK122833 Homo sapi
18	2191.5	73.5	7878	6	AX332512 Sequence
19	2191.5	73.5	7878	9	HSTRE210
20	2188.5	73.4	7971	9	AY143550 Homo sapi
21	2184.5	73.2	8491	9	BSM807865
22	2133.5	71.5	8201	6	I76205
23	2133.5	71.5	8201	6	AX774956 Sequence
24	2133.5	71.5	8201	9	HSTRE213
25	1522	51.0	134506	2	AC146482
26	1522	51.0	180876	9	AC021317
27	1522	51.0	181517	9	AC126327
28	1521	51.0	193646	9	AC131056
29	1518	50.9	146243	2	AC130293
30	1514	50.8	37757	9	AC139091
31	1505	50.5	173053	2	AC129893
32	1501	50.3	140152	9	AC027821
33	1501	50.3	140152	9	AC027821
34	1500.5	50.3	158698	2	AC036181
35	1500.5	50.3	192881	9	AC067923
36	1500	50.3	206040	2	AC132814
37	1497.5	50.2	185449	2	AC139361
38	1494.5	50.1	42073	2	AC144840
39	1493.5	50.1	41266	2	AC139176
40	1493.5	50.1	89147	2	AC146857
41	1487.5	49.9	146243	2	AC130293
42	1469	49.2	123925	9	AC003976
43	1465	49.1	172770	2	AC023133
44	1448.5	48.6	163628	2	AC139367
45	1373.5	46.0	36160	2	AC139089

ALIGNMENTS

RESULT 1
AX775945
LOCUS AX775945 1964 bp mRNA 14near PAT 14-JUL-2003
DEFINITION Sequence 215 from Patent WO03048202.
ACCESSION AX775945
VERSION AX775945.1 GI:32693663
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 215 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source location/Qualifiers
1. 1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 1650
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB1672.1"
/db_xref="GI:32693664"
/translation="MDVVEVAGSWWAQEREDIIKYEKGHRAGLPEDKGPFRSYNN
NVDHLGIYHETELPLLTAREAKQIRREISRKSKWVMDGDWEKYKSSRKLIDRAYKGM

CDS

PMNIRGPMWSVLINIEEMKLNPGRYQIMKEKRSSEHIQRIIDRVSGTLRKHIFR
DRYGTQRELIHLILAYEENYENVEGYCRDLSHIAALFLYLPEEDAFMALVQLLASER
HSLQGFHSNGGTVOGLDQOEHVATSOPTMGHODKDLGQCSPLGCLIRILIDG
ISLGLTRLMDEVYLVEEQALMPITRIAFKVOQKRLITTSRCGPMARFCNRFVDTWAR
DEDYVLKHLIRASMKULTRKQDLPFPAKPEQSSASRPVPSARGKTLCKGDRQAPW
PPARFPRPWSASPPRAPRSSTPCPGGAVREDTYPVGTQGVSPALAQGGPOGQSWRFL
QWNSMRLPTDLIDVEGFWFRHYDFRQSCWVRALSQEDQLAPCWOAEHPAERVRSAFRA
PSTUSDQGTFFRADEQPCAPTSGPCLCGLHLESQPPPGP"

ORIGIN

Alignment Scores:

Pred. No.:	1.12e-103	Length:	1964
Score:	2983.00	Matches:	549
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-071-838-2 (1-549) x AK775945 (1-1964)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
DB	1	ATGAGCGTGTAGAGGTCGCGGCGAGTTGGTGGCACAAGAGCGAGACATCATATG	60
QY	21	LysTyrGluLysGlyHisAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	61	AAATAAGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCCTTAAGCTTTTGA	120
QY	41	SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu	60
DB	121	AGCTACAAACAACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG	180
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
DB	181	ACTGCGCGGAGCGAAGCAATTCGCGGAGATCAGCCGAAGCAAGTGGGTGAT	240
QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
DB	241	ATGCTGGGAGACTGGGAGAATAACAAGACAGAAAGCTCATGATCGAGCTTACAAG	300
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet	120
DB	301	GGAATGCCCATGAAATCCGGGGCCGATGTGTGTCAGTCTCTGAAACATTGAGGAATG	360
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
DB	361	AAATTGAAAAACCCCGAAGATACCATCATGAAGAGAAAGGCAAGGTCATCTCAGG	420
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
DB	421	CACATCCAGCGCATTCGACCGGAGCGTAAGCGGACATTAAAGAACATATATCTTCAAG	480
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
DB	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT	540
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr	200
DB	541	AAACCGGAGGTGGCTACTGACGAGCCTGAGCCACATGCCCTTGTCTCTCTCTAT	600
QY	201	LeuProGluGluAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
DB	601	CTTCCTGAGAGATGCATTCTGGGCACTGTGACGCTGCTGCCACTGAGAGGCACTCC	660
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
DB	661	CTGACGAGGATTCACAGCCCAATGGCGGACCGTCCAGGGCTTCAAGACCAAGAGAG	720
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
DB	721	CATGTGTAGCCACGTCACAAACCAAGACCATGGGTCATCAGGACAAAGAAAGATCTATG	780
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280

DB	781	GGGCACTGTCCCGCTTAGCTGCTCCATCCGATATTGATTGACGGGATCTCTCTGGG	840
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
DB	841	CTCACCTCGCGCTGTGGACCGTATCTGTTAGAAAGCGCAAGCGGTGATGCTGCATA	900
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
DB	901	ACAAGATCCCTTTAAGGTTCAAGACAGAGCCCTCACGAAGACGTCCAGGTGTGCGCG	960
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
DB	961	TGGGCACGTTTTCACACCGGTTCTGTATACCTGGGCCAGGATGAGACACTGTGCTC	1020
QY	341	LysHisLeuAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
DB	1021	AAGCATCTTAGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA	1080
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
DB	1081	GCCAAACCCGACAGAGGTCGTCCGATCCAGGCTGTGCCGCTTACGTCGGGAAG	1140
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
DB	1141	ACCCTGTCAAGGGGAGCAGGCGCCCTCCAGGCCACCAAGCCCGGTTCCCGCGGCC	1200
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla	420
DB	1201	ATTGTGTCAGCTTCCCGGCCACGGGACCTCGTTCTTCCACACCTGCTGTGGGGCT	1260
QY	421	ValArgGluAspThrThrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
DB	1261	GTCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCACGCCGCTGCTCAG	1320
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
DB	1321	GAGGACCTCAGGGTCTTGAGATTCTTCACTGAACTCAATGCCCGCTCCCAACG	1380
QY	461	AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
DB	1381	GACCTGACGTAGAGGGCTTGTGTCGACATTAATGATTACAGACAGAGCTGCTGGTC	1440
QY	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln	500
DB	1441	CGTGCATATCCAGAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTCGCGAG	1500
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
DB	1501	CGGTGAGATCGGCTTTCGCTGCACCCAGCACATGATTCCGACCAAGGACCCCTCAGA	1560
QY	521	AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
DB	1561	GCTAGGACGAACAGCGGTGTCTCCACCTCAGGCGCTTGCCTGTGCGGCTCCACTTG	1620
QY	541	GluSerSerGlnPheProGlyPhe	549
DB	1621	GAAAGTTCTCAGTTCTCTCCAGGCTTC	1647

RESULT 2
AF540953
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

1964 bp mRNA linear PRI 02-OCT-2002
complete cds.
AF540953
AF540953.1 GI:23452664
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Pei, L., Peng, Y., Ling, X.B., van Eynhoven, W.G., Nguyen, K.C.Q.,

TITLE Rubin,M., Hoey,T., Powers,S. and Li,J.
PRC17, a novel oncogene encoding a Rab GTPase-activating protein,
is amplified in prostate cancer
JOURNAL Cancer Res. (2002) In press
REFERENCE 2 (bases 1 to 1964)
AUTHORS Li,J.L., Peng,Y. and Powers,S.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) Genomics, Tularik, Inc, 266 E Pulaski Road,
Suite 1, Greenlawn, NY 11740, USA
FEATURES
source
1. .1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1. .1964
/gene="PRC17"
1. .1650
/gene="PRC17"
/codon_start=1
/product="Rab GTPase-activating protein PRC17"
/protein_id="AAN33117.1"
/db_xref="GI:23452665"
/translation="MDVEVAVAGSWAQRREDIIMKYEKHRAGLPEDKGPFRSYNN
NVDHLGIVHETELPPLTAREAKQIRREISRSKWVDMLGWMEKYSSRKLIDRAYKGM
PMNIRGPMWSVLNLIBEMKLNPEGYQIMKEKGRSSEHIQRIDRVSGTLRKHIFER
DRYGTOKORELHILAYEEYNPEVGYCRDLSHIALFLVLPEDAFWALVOLLASER
HSLQGFHSPNGSTVQGLQDQEHVAVTSQPKTMGHQDKDLGGCSPGLCIRLLIDG
ISLGLTLRLMDVYLVEGEQALMPTRIAFKVQOKRLTTSRCGPWARFCNRFVDTWAR
DEDTVLKLIRASMKLTKRQGLDPPPAKPEQSSASRPVVASRGKTLCKGDRQAPPG
PPAPFPRIWSPAPRARSSTPCGAVREDTYPVGTQGVSPALAQGGPQGSWRF
QWNSMPRLPTDLVEGPMFRHYDRQSCWVRAISQEDQLAPCWQAEHPAERVSAPFAA
PSTSDQGTFRARDEQPCAPTSGLCLHLESSQFPPEP"

ORIGIN

Alignment Scores:

pred. No.: 1.12e-103 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-071-838-2 (1-549) x AF540953 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1 ATGGACGTGTAGAGGTTCGGCGGACGTTGTTGGGCAACAAGCGAGGACATCATTAATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 61 AAATACGAAAGGACACCGAGCTGGGCTGCAGAGCAAGGGGCTTAAGCCTTTTGA 120
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 121 AGCTACAACAACACGTCGATCATTTGGGGAATTGATCATGAGACGGAGCTGCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 181 ACTGCGCGGAGCGGCAACAATTTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACATTGAGGAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGGAAGGCAAGGTCTCATCTGAG 420

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CACATCCAGCGCATCGACCGGACGTAAACGGGACATTAAAGAAAGCATATTCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCGCATATGAGAGATAT 540
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
DB 541 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 601 CTTCCTGAGAGAGATGCAATTCGGGCACTGTGACGCTGTGGCCAGTAGAGGACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
DB 661 CTGCAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAAGAATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 841 CTCACCTGCGCGCTGTGGACGTGTATGTGTAGAAAGCGCAAGCGGTGATGCGGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 901 ACAAGAATCGCCTTAAGTTCAAGCAAGCGCCTCAGAAAGCGTCCAGGTGGCCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 961 TGGGCACTTTTGGCAACCGGTTCTGTGATACCTGGCCAGGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1021 AAGCATCTTAGGCTCTATGAAGAACTTAACAAGAAAGCAGGGGGAACCTGCCACCCCA 1080
QY 361 AlaLysProGlnGlnLysSerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1081 GCCAAACCCGAGCAAGGTCGTGGCAATCCAGGCTGTGCCGCTTACGTCGGGGAAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTGCAAGGGGACAGGACAGGCCCTCCAGGCCACCAAGCCGCTCCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1201 ATTTGGTCAAGCTTCCCCGCCACGGGCACTCTCTTCCACACCTGTCTGTGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGGGAAGACCTTACCTGTGGGCACTCAAGGTGTGCCAGCCCGGCTGCTCAG 1320
QY 441 GlyGlyProGlnLysSerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1321 GAGAGACCTCAGGGTTCCTGAGATTCTGCAGTGAACTCCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1381 GACCTGACGTAGAGGGCCCTGTGTCGCCATTATGATTTCAGACAGAGCTGTGGGT 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGln 500
DB 1441 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGCGAGGCTGAACACCTGCGGAG 1500

QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1501 CGGTGAGATCGCTTCCTGCTGCACCCAGCACTGATCCGACCAAGGCAACCCCTTCA 1560
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 1560
Db 1561 GCTAGGAGACGAACAGCGGTGTGCTCCCACTCAGGGCTTGCTCTGCGGCTCCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1621 GAAAGTCTCAGTTCCTCCAGGCTTC 1647

RESULT 3
AX775947
LOCUS AX775947 1964 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 217 from Patent WO03048202.
ACCESSION AX775947
VERSION AX775947.1 GI:32693665
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 217 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source location/Qualifiers
1. 1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 1650
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11673.1"
/db_xref="GI:32693666"
/translation="MDVVEVAGSGWMAQEREDIMKYEKHRAGLPEDKGPFRSYNN
NVDHGLIVHETELPPLTAREAKQIRREISRSKWVMDLGDWEKYSRKLLIDRAYKGM
PMNIRGPMWSVLNIEEMKLNKRGYQIMKEKGRSSEHIQIRDRDVSGLTRKHIFER
DRYGTOKORELHILLAYEENDEVCYCRDLSHIALFLYLPEDAFWALVQLASER
HSLQGFHSNNGGTVOQLDQOEHVATISQPKTMGHQDKDLGQCSPLGLIRILIDG
ISLGLTLRLMDVYLVEGEALMPITRIAFKVVQKRLTSTKSRGPMWAFNRPVDTWAR
DEDTVLKLRLASMKLTLKQGLDLPAPKEGSSASRPVPSRGGKTLCKGDRQAPG
PPARFPRPIWSASPPRARSSTPCPGAVREDTYPVGTGVSPPALAOGSGSWRFL
QWNSMPRLPTDLVEGPMFRHYDFRQSCWVRAISQEDQLAPCWOAEHPAERVRSFAFA
PSTSDQGTFRARDEQPCAPTSGLCLGLHLESSQPPGF"

Alignment Scores:
Pred. No.: 2.23e-103 Length: 1964
Score: 2975.00 Matches: 548
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.73% Indels: 0
Gaps: 0

US-10-071-838-2 (1-549) x AX775947 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGGACGTGTAGAGTCGCGGCACTGTGGGCACAAGAGCGAGAGACATCATTAATG 60
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTCA 120
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGlyThrGlnLeuProProLeu 60
Db 121 AGCTACAAACAACGTCGATCATTTGGGATGTATGATGAGACGAGCTGCTCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80

Db 181 ACTGCGCGGAGCGCAAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
QY 81 MetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTCATGATCGAGCTTACAAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGluMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGACATTAAGGAAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAAACCCCGGAATACCAAGATCATGAAGAGAGAGGCAAGCTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGAGCTTAAGCGGACATTAAGAGACATATTCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGlnLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
Db 481 GATCGATACGGAACCAAGCAGCGGGAATCACTCAATCCTCTGCGCATATGAGAGATAT 540
QY 181 AsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
Db 541 AACCCGGAGGTGGCTACTGTCAGGGACCTGAGCCACATCGCGCCTGTCTCTCTCTAT 600
QY 201 LeuProGlnGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCTGAGAGAGATGATCTTGCGCACTGTGTGACGTCTGCGCAGTGAGGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspDlnGlnGlu 240
Db 661 CTGCAAGGATTTCAAGCCCAATGCGGAGCCGTCCAGGGCTCCAGACCAACAGAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 721 CATGTGTAGCCACGTCAACCAAGACCATGGGCGCATCAGAGACAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 841 CTACCCCTGCGCTGTGGACGTGATCTGTAGAAAGCGAACAGCGGTGATGCCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAAGAAATCGCCTTAAAGTTACAGCAGAAGCGCTCAGCAAGACGTCCAGGTGTGCGCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 961 TGGGCACTTTTTCACACCGGTTGTTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1021 AAGCATCTTAGGCTCTTATGAGAAACTAACAGAAAGCAGGGGACCTGCAACCCCA 1080
QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGCGCTGTGCGGCTTACAGTGGCGGAAG 1140
QY 381 ThrLeuCybLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1141 ACCCTTCGCAAGGGGACAGGAGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
Db 1201 ATTGTGTCAGCTTCCCGCCAGCGGACCTGTTCTTCCACACCTGTCTGTGGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

Db 1261 GTCCGGGAAGACACTACCTGTGTGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
Qy 441 GYGLYPRGGLNGLYSERTRPARGPHELEUNGINTPANSERMETPROARGLEUPROTHR 460
Db 1321 GGAGGACCTCAGGGTCTCTGAGATTCCTGACGTGAATCCATGCCCGCTCCCAACG 1380
Qy 461 AAPLEUASPVALGLUGLYPROTRPHEARGHISTYRASPAPHEARGINSERCYSTRVAL 480
Db 1381 GACCTGACGTAGAGGGCCCTTGCTCCGCCATTGATTTGACAGACAGAGCTGGGTG 1440
Qy 481 ARGALAILESERGLNGLUASPGINLEUALAPROCYSTRGINALAGLUHISPROALAGLU 500
Db 1441 CGTGCCATATCCCAAGAGACCACTGGCCCCCTGTGCGCAGGCTGAACACCTCGCGAG 1500
Qy 501 ARGVALARGSERALAPHEALAPROSERTHRASPSEARSPGLNGLYTHRPROPHARG 520
Db 1501 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGCGACCCCTTCAGA 1560
Qy 521 ALAARGASPGIUGINPROCYALAPROTHRSERGLYPROCYLEUCYSGLYLEUHLISLEU 540
Db 1561 GCTAGGAGCAGAACAGCCGCTGTGCTCCACCTCAGGGCCTTGCTGTGCGGCTCCACTTG 1620
Qy 541 GLUSERSERGLNPHETPROGLYPHE 549
Db 1621 GAAAGTCTCTAGTTCCTCCAGGCTTC 1647
RESULT 4
BC071680 2090 bp mRNA linear PRI 30-JUN-2004
LOCUS BC071680
DEFINITION Homo sapiens TBC1 domain family, member 3, mRNA (cDNA clone
MGC:87891 IMAGE:5165385), complete cds.
ACCESSION BC071680
VERSION BC071680.1 GI:47939522
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2090)
AUTHORS Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Helel,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Mortley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huljk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,B.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,B.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2090)
AUTHORS Strausberg,R.
TITL Direct Submission
JOURNAL Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@pax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 166 Row: g Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984.
FEATURES
source
1..2090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87891 IMAGE:5165385"
/issue_type="Brain, adult medulla"
/clone_id="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2090
/gene="TBC1D3"
/note="synonym: PRC17"
/db_xref="locusID:84218"
/db_xref="MIM:607741"
101..1750
/gene="TBC1D3"
/codon_start=1
/product="TBC1 domain family, member 3"
/protein_id="AAH71680.1"
/db_xref="GI:47939523"
/db_xref="LOCUSID:84218"
/db_xref="MIM:607741"
/db_xref="MIM:607741"
/translation="MDVEVAGSWWAQEREDIMKYEGHAGLPEDKPKPFRSYNN
NVDLGIIVHETELPPLTAREAKQIREISRSKSMVMDLGMEEKYSSRKLIDRAYKM
PMNIRGPMWSVLINIBEMKLNPGRYQIMKEGKRSSEHQRIDRISGTLRKMPFR
DRYGTQKRELHILLVLAEEYNPEVGYCRDLSHIALFLYLPEDEAFWALVOLASER
HSLOGFHSFNGTIVQLODOEQHEVAVTSQPTMGHODKDLGCGSPGLIRILIDG
ISLGLTRIMDVYLVGEQALMPITRIAEKVQQRLLTKTSRCGPWAFNRFVDTAR
DEDTVLKHILASMKLTKRQGLDLPBAKEGSSASRPVPSRGKTLCKGDRPADPG
PPARFPRIWSPAPRAPRSSTPCGAVREDTYPVGTQVSPALAOQGGPOGSMREL
QNSMPRLPTDLVDVEGFWFRHYDRROSCWVRASIQEDQLAPCWAHPAERVRSAFPA
PSTDSDQTPFRARDEQCAFTSPCLCGLHLESSQFPFG"

ORIGIN
Alignment Scores:
Pred. No.: 3.35e-103 Length: 2090
Score: 2971.00 Matches: 546
Percent Similarity: 99.82% Conservative: 2
Best local Similarity: 99.45% Mismatches: 1
Query Match: 99.60% Indels: 0
DB: 9 Gaps: 0
US-10-071-838-2 (1-549) x BC071680 (1-2090)
Qy 1 Metaspvalalgluvalalaglysertrtpalaglgluarggluaspillelmet 20
Db 101 ATGGAAGTGTAGAGGTCCGGGCACTGTGCGGCACAGAGCGAGAGACATCATATG 160
Qy 21 LysTYRGLULYGLYHISARGALAGLYLEUPROGLUASPLYSGLYPROLYSPROPHARG 40
Db 161 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAGAGGCGCTTAAGCCTTTTCA 220
Qy 41 SerTYRAsnaAsnaValaspHisleuGLYILEVALHISGLUthrgluLeuProProleu 60
Db 221 AGCTACACACACACATTCGATCATTTGGGATTGTACATGAGAGGAGCTGCTCTCTG 280
Qy 61 ThrAlaARGLUAlaLYSGINleargARGLUileSerargLYSerLYSTrpValasp 80
Db 281 ACTGCGCGGAGCGCAACAAATTCCGCGGAGATCAGCCGAAAGCAAGTGCGTGAT 340

QY 81 MetLeuGlyAspTrpGluLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 341 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATGATCGAGCGTACAAAG 400
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 401 GGAATGCCCATGAAACATCCGGGGCCCGATGTGTCACTCTCTGAAACATTGAGGAAATG 460
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysArgSerSerGlu 140
Db 461 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAGGCAAGAGGTATCTGAG 520
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 521 CACATCCAGCCGATCGACCGGACATTAAGCGGACATTAAGGAAAGCATATGTTCTTCAGG 580
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 581 GATCGATACGGAAACCAAGCAGCGGAACTACTCCACATCTCTGCGCATATGAGAGTAT 640
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 641 AACC CGAGGTGGGCTACTGACGAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 700
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 701 CTTCTGAGAGAGATGATTCATTCGTGGCAGCTGTGCTGCGCATGAGAGGCACTCC 760
QY 221 LeuGlnGlyPheHisSerProAsnGlyTyrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 761 CTGACAGGATTTCAACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 820
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 821 CATGTGTAGCCACGTCACAAACCAAGCATTGCGGATCAGGACAAAGAAATCTATGT 880
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 881 GGGCAGTGTCCCCGTTAGGCTGAGCGTCTCATCCGATATGATGACGGGATCTCTCTGGG 940
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 941 CTCACCTCGGCTGTGGGACGTGATCTGTAGAGGCGGAACAGCGCTTGATGCCGAT 1000
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1001 ACAAGATCGCCTTTAAGTTACAGCAAGCGCTCACAGAACGTCACAGTGTGCGCCG 1060
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1061 TGGGACGCTTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGATGAGGACACTGTGCTC 1120
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1121 AAGCATCTTAGGGCCTTATGAAAGAACTAACAAAGAGGAGGAGCACTGCCACCCCA 1180
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
Db 1181 GCCAAACCCGAGCAAGGCTGTGGCATCAGGCTGTGCGGCTTCACTGCGGGGAG 1240
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1241 ACCCTCTGCAAGGGGACAGGAGGCTCCAGGCCACCAAGCCGTTCCCGCGGCC 1300
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1301 ATTTGTACGCTTCCCGCCAGCGGACCTGTTCTTCCACACCTGTCTGTGTGGGCT 1360
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1361 GTCCGGAGACACCTACCTGTGGGCACTCAGGGGTGTGCCACGCCGCTGCTCAG 1420

QY 441 GLYGLYProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1421 GGAGGACCTCAGGGTCTCTGAGATTCCTGACGTGAACTCCATGCCCGCTCCCAACG 1480
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1481 GACCTGACGTAGAGGGCCCTTGCTTCCCGCATATGATTTCAGACAGAGCTGTGGTTC 1540
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1541 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGAG 1600
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1601 CGGTGAGATCGGCTTTCCTGACCCACCACTGATTCGACCAAGGCAACCCCTTCAGA 1660
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1661 GCTAGGAGACGACAGCAGTGTGCTCCACCTCAGGGGCTTGCTGCGGCTCCACTTG 1720
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1721 GAAAGTTCTCAGTTCCTCCAGGCTTC 1747
RESULT 5
AX086847
LOCUS AX086847 2072 bp DNA linear PAT 09-MAR-2001
DEFINITION Sequence 799 from Patent WO0112659.
ACCESSION AX086847
VERSION AX086847.1 GI:13276047
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Wiemann, S., Human dna sequences Patent: WO 0112659-A 799 22-FEB-2001;
AUTHORS German Human Genome Project (DE)
TITLE Journal location/Qualifiers
JOURNAL 1. .2072
FEATURES
source /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.63e-103 Length: 2072
Score: 2963.00 Matches: 545
Percent Similarity: 99.64% Conservative: 2
Best Local Similarity: 99.27% Mismatches: 2
Query Match: 99.33% Indels: 0
DB: 6 Gaps: 0
US-10-071-838-2 (1-549) x AX086847 (1-2072)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 99 ATGACGTGTAGAGGTGCGGGCAGTGTGTGGGCACAAAGAGCGAAGACATCATATATG 158
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 159 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 218
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 219 ACCTACAAACAACAGCTGATCATTTGGGATTTGTACATGACAGGAGCTGCTCTCTG 278
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 279 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAAGCGAAGAGCAAGTGGGTGAT 338

QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
Db	339	ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATGATCGAGCGTACAG	398
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValIleuLeuAsnIleGluMet	120
Db	399	GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTCTGAACACTGAGAAATG	458
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysArgSerSerGlu	140
Db	459	AAGTGAATAACCCCGGAACATACCAAGATCATGAAAGAAAGGCAAGAAGTCATCTGAG	518
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	519	CACATCCAGCGCATGCAACCGGACGTAAACGGGACATTAAAGAACATATATTCTTCAGG	578
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	579	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAC	638
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr	200
Db	639	AACCCGAGGTGGGCTACTGCAAGGAACTGAGCCACATCGCCGCTTGTCTCTCTAT	698
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	699	CTTCCTGAGAGATGCAATTCTGGGCACTGTGTGACGTGCTGGCCAGTGAGAGGCACTCC	758
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	759	CTGCAGGGAATTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG	818
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	819	CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAAGACAAGAAAGATCTATGT	878
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	879	GGGCAGTGTCCCGTTAGGCTGCCCTCATCCGATATTGATTGACGGGATCTCTCTGGG	938
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
Db	939	CTCACCTCGCGCTGTGGACGTGTATCTGTAGAAGCGAACAGCGCTGATGCCGATA	998
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	999	ACAAGAATCGCCTTTAAGTTCAAGCAGAGCGCCTCACGAAGACGTCCAGGTGTGGCCCG	1058
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1059	TGGGCACGTTTTTGCACCGGTTCTGTATACCTGGGCCAGGAGATGAGACACTGTGTC	1118
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1119	AAGCATCTTAGGCGCTCTATGAAGAACTAACAAAGAAAGAGGGGAGCTGCCACCCCA	1178
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
Db	1179	GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCGCTGTGCGGCTTCACTGCGGGAAG	1238
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1239	ACCCTCTGCAAGGGGAGCAGGCCCCCTCCAGGCCCAACGCGGTTCCCGCGGCC	1298
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1299	ATTTGGTCAGCTTCCCGGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT	1358
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1359	GTCCGGGAAGACACCTTACCTGTGGGCACTCAGGGGTGTGCCACGCCGGCCCTGGCTCAG	1418
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460

DB 1419 GGAGGACCTCAGGTTCTCGAGATTCCTGCAGTGGAACTCCATGCCCCGCTCCCAACG 1478

QY 461 ASPLeuASPValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480

DB 1479 GACCTGGACCGTAGAGGGCCCTTGTTCCGCCATTATGATTTTCAGACAGAGCTGCTGGGTC 1538

QY 481 ARGAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500

DB 1539 CGTGCCATATCCCAAGAGACCAGCTGAGCCCTCTGCGCAGGCTGAACACACCCTGGCAG 1598

QY 501 ARGValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520

DB 1599 CGGGTGAGATCGGCTTTTCGCTGCACCCAGCACTGATTCCGACCAGGGCACCCCCTTCAGA 1658

QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540

DB 1659 GCTAGGAGCAACAGCAGTGTGCTCCCACTCAAGGCGCTTGCTGCGGCGCTCCACTTG 1718

QY 541 GluSerSerGlnPheProProGlyPhe 549

DB 1719 GAAAGTTCACGTTCCCTCCAGGCTTC 1745

RESULT 6				
HSMB01828	HSMB01828	2072 bp	mRNA	linear
LOCUS	Homo sapiens mRNA; cDNA DKFZp434P2235 (from clone DKFZp434P2235);			
DEFINITION	complete cds.			
ACCESSION	AL136860			
VERSION	AL136860.1	GI:12053224		
KEYWORDS				
SOURCE	Homo sapiens (human)			

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2072)
REFERENCE	Poustka, A., Klein, M., Mewes, H.-W., Gassenhuber, J. and Wiemann, S.
AUTHORS	Direct Submission
JOURNAL	Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp434P2235) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcencentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES	Location/Qualifiers
source	1. .2072

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434P2235"
/db_xref="taxon:9606"
/map="17"
/clone="DKFZp434P2235"
/tissue_type="testis"
/clone_lib="434 (synonym: htess). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2072
/gene="DKFZp434P2235"
99..1748
CDS

```

```

/gene="DNKFZp434P2235"
/note="strong similarity to oncogene 1 (tre-2 locus)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66794.1"
/db_xref="GI:12053225"
/db_xref="TrEMBL:Q9H0B9"
/translation="MDVVEVAGSWMQGERDIIIMKYEKGRAGLPEDKPKPFRSYNN

```

NVDHLGIVHETELPDLTAREAKOIRREISRSKWVMDLGDWEKYSRKLTIDRAKGM
PMNIRGPMWSVYLINTEEMKLNKNGRYQIMKEKSKSSEHIORIDRDVSGTLARKHIFER
DRYGTOKORELHILLAYEBYNPEVGYCRDLISHIALFLYLPEBDAFWALVOLLASER
HSLQGFHSNPGSTVOGLODOEHVVAISOPTWGHODKDLCCGCSPLGCIIRILIDG
ISLGLTLRLMDVYLVEGEQALMPITRIAFYVQOKRLTKTSRCGPWARFCNRFDVTWAR
DEDTVLKLIRASMKLTTRKGDLPPEAKPEQSSASRPVPSRGKTLCKGDROAPPG
PPARFPRIWASPPRARSSTPCPGAVREDTYPVGTOGVSPALAQGGPQGSWRFL
QWNSMPLRLPTDLVEGPFRRHYDFRQSCWVAISQEDOLAPCWOAEHPRERVRSAPFA
PSTDSOGTFPRARDEQCAPTSGPCLCGLHLESSQFPDPGF"
2040. . 2045
polya_signal
/gene="DKFZp434P2235"
2062
polya_site
/gene="DKFZp434P2235"

Alignment Scores:
Pred. No.: 6.63e-103 Length: 2072
Score: 2963.00 Matches: 545
Percent Similarity: 99.64% Conservative: 2
Best Local Similarity: 99.27% Mismatches: 2
Query Match: 99.33% Indels: 0
DB: 9 Gaps: 0

US-10-071-838-2 (1-549) x HSM801828 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 99 ATGACCGTGTAGAGGTCCGGGGCAGTTGTTGGGCACAAAGCGAGACATCATTAATG 158
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 159 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTCA 218
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 219 AGCTACACACACACACGTCGATCATTTGGGATTGTTACATGAGACGGAGCTGCTCTCTG 278
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 279 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCCGAAAGACMAAGTGGTGGAT 338
QY 81 MetLeuGlyAspTrpGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 339 ATGCTGGAGAGACTGGAGAAATACMAAGACAGAGAAAGCTCATGATCGAGCTACAAAG 398
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 399 GGAATGCCCATGAACATCCGGGGCCCATGTGTGTCAGTCTCTCTGAACACTGAGGAATG 458
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 459 AAGTTGAAAAACCCCGGAGATACCATCATGAGAGAGAGGCAAGATCATCTGAG 518
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 519 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAGGAAGCATATATTCTTCAGG 578
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 579 GATCGATACGGAACCAAGCAGCGGACTACTCCACATCCTCTGACATATGAGAGATAC 638
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db 639 AACCCGAGAGTGGGCTACTGACGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT 698
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 699 CTTCCTGAGAGGATGATTCGTGGCACTGTGCACTGTGGCAGTGAAGGACTCC 758
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 759 CTGCAGGGAATTTCACAGCCCAATGGCGGAGCGTCCAGGGGCTCCAAAGCAACAGAG 818

QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 819 CATGTGTAGCCACCGTCACCAACCCAGACCATATGGGGCATCAGACAAAGAAAGATCTATGT 878
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 879 GGGCAGTGTTCCTCCGTAGGCTGCTCCTCATCCGGATATTGATTGACGGGATCTCTCCGGG 938
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnGlnAlaLeuMetProIle 300
Db 939 CTCACCCCTGCCCTGTGGAGGTGTATCTGTTAGAAAGCGCAAGCGCTGATGCCGATA 998
QY 301 ThrArgIleAlaPheLeuValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 999 ACAAGATCGCCTTTAAGGTTCAAGAGACCGCTCAGAAAGACGTCCAGGTGTGCGCCG 1058
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1059 TGGGCAGTTTGTGCAACCGGTTCTTGATACCTGGGCCAGGGATAGAGACACTGTGCTC 1118
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1119 AAGCATCTTAGGCGCTCTATAGAAACCTAACAAAGAAAGAGGGGACCTGCCACCCCA 1178
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1179 GCCAAACCCGAGCAAGGTCGTGGCATCCAGCGCTGCGGGCTTCACTGTGGCGGAG 1238
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1239 ACCCTCTGCAAGGGGAGCAGGACGGCCCTCCAGGCCACAGCCCGGTCCCGCGGCC 1298
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1299 ATTTGTCAGCTTCCCGCCACGGGCACTCGTCTTCCACACCTGTCTGTGGGGCT 1358
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1359 GTCCGGAGAACACCTACCTGTGTGGCACATCAGGGGTGTGCCAGCCCGCGCTGGCTCAG 1418
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1419 GGAGAGACTCAGGTTCTGTGAGATTCTGCACTGAGACTCCATGCCCCCTCCCAACG 1478
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1479 GACCTGGAAGTGAAGGCGCTTGTGTTCCGCATATGATTTCAGACAGAGACTGTGGCTC 1538
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1539 CGTGCCATATCCAGAGAGACCACTGGCCCTGCTGGCAGGCTGAACACCTCGCGAG 1598
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1599 CGGTGAGATCGGCTTTCGTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1658
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1659 GCTAGGAGCAAGACAGAGTGTCTCCCACTCAGGGCTTGCTCTGCGGCTCCACTTG 1718
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1719 GAAAGTTCTCAGTTCTCTCCAGGCTTC 1745

RESULT 7
BC075809
LOCUS
DEFINITION
Homo sapiens similar to TBC1 domain family, member 3, centromeric,
mRNA (cdna clone MGC:87892 IMAGE:5170064), complete cds.
ACCESSION
BC075809
VERSION
BC075809.1 GI:49902676
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2166)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalius, D.E., Scherch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 12477932
PUBMED 2 (bases 1 to 2166)
REFERENCE Director MGC Project.
AUTHORS Direct Submission
TITLE Submitted (06-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 166 Row: h Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.
FEATURES
Source
1..2166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87892 IMAGE:5170064"
/issue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2166
/gene="LOC414059"
/db_xref="LocustID:414059"
147..1796
/gene="LOC414059"
/codon_start=1
/product="centromeric TBC1 domain family member 3-like"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="LOCustID:414059"
/translation="MDVVEVAGSWAEDREDIMKYEKGHAGLPEDKGPFRSYN
NVDHLGIWHTLPLTLARAKQIRREISRSKSWVDMLGDWEKYSRKLLIDRAYGM

PMNIRGPWMSVLLNIEEMKLNKPGRYQIMKEKGRSSEHIQIDRDISGTLRRHMF
DRYGTQKRELHILLAYEENPVEVYCRDLSHIALFLYLPEEDAFWALVOLASER
ISLOGFHSFNGCTVOGLQDQDQEHVATIASQSKTMHODKCGQCSPLGLIRLIDR
ISLGLTLRLMDVYLVEGEQALMPLTRIAFKVQQRLLTKTSRCGFWARFCNRPVDTWAR
DDDTVLKHLRASMKKLTRKQGLDLPBPAKEQSSASRPVPASRGKTLCKGDRQAPG
PPARFPRIWASPPRAPRSSTPCGAVRBDTPVGTQVPSPALAQGGPOGSMRFL
QNSMRPLPTDLVEGFWFRHYDRQSCWVRASIQEDQLPCWQAEHPAERVRSAPFA
PGTSDQGTPEPRARDEQQCAPTSBCLGLHLESSQFPFGP"
ORIGIN
Alignment Scores:
Pred. No.: 3.27e-102 Length: 2166
Score: 2945.00 Matches: 541
Percent Similarity: 99.27% Conservative: 4
Best Local Similarity: 98.54% Mismatches: 4
Query Match: 98.73% Indels: 0
DB: 9 Gaps: 0
US-10-071-838-2 (1-549) x BC075809 (1-2166)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 147 ATGAACGTGTAGAGGTCCGGGTAGTGTGTGGGCACAAGAGCGAGGACATCATATG 206
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 207 AAATACGAAAGGGACACCGAGCTGGGCTGCCAGAGACAGAGGAGGGCTTAAGCCTTTTCCA 266
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 267 AGCTACACACACACGCTGATCATTTGGGGATTGTACATGAGACGGAGCTGCTCTCTG 326
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 327 ACTGCGCGGAGCGCAACCAATTCCGGGAGATCAGCCGAAAGCAAGTGGTGAT 386
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 387 ATGCTGGAGACTGGGAGAAATACAAAGAGACAGAAAGCTCATGATCGACGTACAAG 446
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 447 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAATG 506
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 507 AAGTTGAAAACCCCGAAGATACCATGATGAAGAGAGGCAAGGTCATCTAG 566
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 567 CACATCCAGCGCATGACCGGACATTAAGCGGACATTAAAGAGCATATGTTCTTCAG 626
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLys 180
DB 627 GATCGATACGGAACCAAGCAGCGGAACCTACCATCTCTCTGACATATGAGAGTAT 686
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLys 200
DB 687 AACCCGAGGTGGGTACTGACGAGCACTGAGCCACATCGCGCTGTCTCTCTAT 746
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 747 CTTCTGAGGAGATGATCTTCTGGGCACTGTGTCAGCTGCTGGCCAGTAGAGGCACTCC 806
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
DB 807 CTGCAAGGATTTCACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAGAG 866
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 260
DB 867 CATGTGTAGCCACGTACCAATCCAAAGACATGGGGCATCAGGACAAAGAAATCATGT 926
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280

Qy	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	962	CTTCCTGAGAGAGATGCATCTGGGCACTGGTGACAGCTGTCGGCAGAGAGACTCC	1021
Qy	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	1022	CTGCAGGGATTTCACAGCCCAATATGCGGGACCGTCCAGGGGCTCCAGACCAACAGAG	1081
Qy	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuGly	260
Db	1082	CATGTGTAGCCACGTCAACACCAGACCATGGGGCATCAGGACAGAAAGATCTATGT	1141
Qy	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	1142	GGGAGATGTTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG	1201
Qy	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle	300
Db	1202	CTCACCTGCGCCTGTGGGACGTGTATCTGGTAGAAGCGAACAGCGCTGTATGCCGATA	1261
Qy	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	1262	ACAAGAATCGCCTTAAAGTTACAGACAGAGCCCTCACGAAGCTCCAGGTGTGGCCG	1321
Qy	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1322	TGGGCACGTTTGTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC	1381
Qy	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1382	AAGCATCTTAGGGCTTATGAAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCCA	1441
Qy	361	AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys	380
Db	1442	GCCAAACCCGAGCAAGGCTCGTCGCATCCAGGCTGTGCGGCTTCACTGCGGGAG	1501
Qy	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1502	ACCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGC-----CCA	1540
Qy	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1541	TT-TGTCAGCTTCCCCGCCACGGGCACTCTGTTCTTCCACACCTCTGCTGTTGGGCT	1599
Qy	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1600	GTCCGGAGAGACACTTACCTGTGGGACATCAGGTTGTGCCAGCCGCTGTGGCTCAG	1659
Qy	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	1660	GGAGACCTCAGGTTCTGTGAGATTCTGTGAGAACTCCATGCCCGCTCCCAACG	1719
Qy	461	AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
Db	1720	GACCTGAGCTAGAGGGCCCTTGGTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC	1779
Qy	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln	500
Db	1780	CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG	1839
Qy	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
Db	1840	CGGGTAGATCGGCTTTGCTGTGCACCCAGACATGATTCCGACCAAGGCAACCCCTTCAGA	1899
Qy	521	AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
Db	1900	GCTAGGAGCAACAGCGGTGTGCTCCACACTCAAGGCTTGTGCTGTGCGCTCCACTTG	1959
Qy	541	GluSerSerGlnPheProProGlyPhe	549
Db	1960	GAAAGTTCTCAGTTCCCTCCAGGCTTC	1986

AX775941	LOCUS	AX775941	2304 bp	mRNA	linear	PAT 14-JUL-2003
DEFINITION	Sequence 211 from Patent WO03048202.					
ACCESSION	AX775941					
VERSION	AX775941.1	GI:32693659				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 Matsuda, A. and Muramatsu, S.					
AUTHORS	NP-KB activating gene					
TITLE	Patent: WO 03048202-A 211 12-JUN-2003;					
JOURNAL	Asahi Kasei Kabushiki Kaisha (JP)					
FEATURES	Location/Qualifiers					
source	1. .2304					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	362. .1732					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAE11670.1"					
	/db_xref="GI:32693660"					
	/translation="MDVVEVAGSWWAQERDIIIMKYEKGHRAGLPEDKGPKPRSYNN					
	NVDHLGIHTEBELPPLTAREAKQIRREISRSKSWMDMGDEKYKSRKLIIDRAYKM					
	PMNIRGPWWSVLNTEEMKLNPNRGYQIMKEKGKRSSEHIQRIDRVSGLRKHIFER					
	DRYGTQORELHILILAYEYNPEVGYCRDLSHIALFLYLPEBDAFWALVQLASER					
	HSLOGFHSPNGTVOGLDDQEHVATSQPKTMGHODKDLGQCSPLGCLIRLLIDG					
	ISLGLTRLMDEVYLVGEQALMPITRIAFKVOQKRLTTSRCGPWFARFCNRPVDTWAR					
	DEDTVLKHLRASMKCLTRKQGLDQPAKPEQSSASRPVPSRGKTLCKGDRQAPPG					
	PFGLPRHGHVLPHPVVLVLSGKPTPTLMALRVCPARFPLREDLRVPDSCSGTIPCBA					
	SQRTWT"					
ORIGIN						
Alignment Scores:						
Pred. No.:	1.92e-100	Length:	2304			
Score:	2898.50	Matches:	539			
Percent Similarity:	98.18%	Conservative:	0			
Best Local Similarity:	98.18%	Mismatches:	3			
Query Match:	97.17%	Indels:	8			
DB:	6	Gaps:	1			
US-10-071-838-2 (1-549) x AX775941 (1-2304)						
Qy	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20			
Db	362	ATGACGTGTAGAGGTGCGGGCAGTTGTGGGACCAAGAGCGGAGACATCATATTATG	421			
Qy	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40			
Db	422	AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTGA	481			
Qy	41	SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu	60			
Db	482	AGCTACACACACACAGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCTCTTG	541			
Qy	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80			
Db	542	ACTGCGCGGAGGCGAAGCAATTCCGCGGAGATCAGCCGAAAGACCAAGTGGTGAT	601			
Qy	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100			
Db	602	ATGCTGGAGACTGGGAGAAATACAAAGCAGACAAAGCTCATGATCGAGCGTACAAG	661			
Qy	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet	120			
Db	662	GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTTGAACACTGAGGAATG	721			
Qy	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140			
Db	722	AAGTTGAAAAACCCCGAAGATACCAATCATGAAGAGAAAGGCAAGAGTCACTCTGAG	781			

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 782 CACATCCAGCGCATCGACCGGGACGTAAAGCGGACATTAAAGAACATATATTCTTCAGG 841
QY 161 AspArgIYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 842 GATCGATACGGAACCAAGACCGGGAACTACTCCACATCTCTCTGGCATATAGAGAGTAT 901
QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTYr 200
Db 902 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCCTTGTTCTCTCTAT 961
QY 201 LeuProGluGluAspAlaPheThrPalaleuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 962 CTTCTTGAGGAGATGCATTTCTGGGCACTGTGCACTGTGCGCAAGTGAAGGCACTCC 1021
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyTYrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1022 CTGACGGGATTTACAGCCCAATGCGGGACCGTCCAGGGGCTCCAGACCAACAGAGAG 1081
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 1082 CATGTGGTAGCCACGTTCACCAACCAAGACCATGGGGCATCAGACCAAGAAAGATCTATGT 1141
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1142 GGGCAGTGTTCCTCCCTTAAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 1201
QY 281 LeuThrLeuArgLeuLeuTrpAspValTYrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 1202 CTCACCCCTGCGCTGTGGGACGTGTATCGGTAGAGGCGCAACAGGCGCTTGATGCCGATA 1261
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1262 ACAAGATTCGCTTTAAGTTCAAGCAAGACGCTTCAGAGACGTCCAGGTGTGCGCG 1321
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1322 TGGGCACTTTTTCACACCGGCTGTTGATACCTGGGCGCAGGATGAGGACACTGTGCTC 1381
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1382 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGACAGGGGAGCTGCAACCCCA 1441
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLys 380
Db 1442 GCCAAACCCGAGCAAGGGTGTGCGGATCCAGGCTGTGCGGCTTCACTGCGGGAG 1501
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1502 ACCCTTGTCAAGGGGACAGGCGCCCTCCAGGC-----CCA 1540
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1541 TT-TGGTCAGCTTCCCGCCACGGGCACTGTTCTTCCACACCTGTCTGTTGGGCT 1599
QY 421 ValArgGluAspThrTYrProValGlyThrGlnGlyValIleProSerProAlaLeuAlaGln 440
Db 1600 GTCCGGGAAGACACCTTCTGTGGGCACTCAGGTTGTGCCACGCCGCTGCTCAG 1659
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1660 GGAGGACCTCAGGTTCTGTGAGATTCTGTGAGTAACCTCAATGCCCGCTCCCAACG 1719
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTYrAspPheArgGlnSerCysTrpVal 480
Db 1720 GACCTGACGTAGAGGGCTTGTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1779
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 1780 CGTGCCATATCCAGAGGACACAGCTGCGCCCTGCTGCGAGGCTGAAACACCTCGCGAG 1839

QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1840 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1899
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1900 GCTAGGGAGCAAGACCGGTGTGCTTCCACCTCAGGGGCTTGTCTGTGCGGCTCACTTG 1959
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1960 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1986
RESULT 10
BC033670
LOCUS
DEFINITION
BC033670 Homo sapiens similar to TBC1 domain family, member 3, telomeric,
ACCESSION
BC033670 mRNA (cdna clone MGC:44903 IMAGE:5167394), complete cds.
VERSION
BC033670.1 GI:45708750
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2575)
AUTHORS
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Rahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2575)
Strausberg,R.
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgonov,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 68 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.

FEATURES

Source
1. .2575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44903 IMAGE:5167394"
/issue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2575
/gene="MGC44903"
/db_xref="locusid:414060"
78. .1208
/gene="MGC44903"
/product="similar to TBC1 domain family, member 3"
/protein_id="AAH33670.1"
/db_xref="gi:45708751"
/db_xref="locusid:414060"
/translation="MDVVEVAVASMAQEREDIMKYEKHRAGLPEDKPKPRSYNNVDHLGIHETELPLTAREAKQIREISRSKMDMLGDMKEYSSRKLDRAVGM
PMNIRGPMWSVLNIEEMKLNPGRYQIMKEKRSSEHIQRIDRISGTLRKMPFR
DRYGTKORELHILLAYEBYNPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER
HSLQGFHSPNGTVQLQDOQEHVAVTSQSKTMGHQDKDLGGQCSPLGLIRILIDG
VRHRETIAQPSLPCSALLPQGLAHSQPTGSGSGSPDTQAAPSAQEGSSQGNMG
GSGPASWADWARTLRLERAQSLKPMASPSLQEPHPSNECP"

ORIGIN

Alignment Scores:

Pred. No.: 1.54e-98 Length: 2575
Score: 2849.00 Matches: 543
Percent Similarity: 77.41% Conservative: 2
Best Local Similarity: 77.13% Mismatches: 4
Query Match: 95.51% Indels: 155
DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x BC033670 (1-2575)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluValArgGluAspIleIleMet 20
DB 78 ATGACGCTGTAGAGCTCGCGGGTAGTGTGGGCAACAAGCGAGACATCATATATG 137
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 138 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAACAAGGGCTTAAGCCTTTTCA 197
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 198 AGCTACAAACAACGTCGATCATTGGGATGTACATGAGACGAGCTGCCCTCTTG 257
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 258 ACTGCGCGGAGCGAACAATTCCGCGGAGATCACGCCGAAAGACAAGTGGTGAT 317
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 318 ATGCTGGGAGACTGGGAGAATACAAAAGCAGCAAAAGCTCATAGTCGAGCTACAAG 377
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 378 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAATG 437
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 438 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGGAGGCAAGAGGTCTCTGAG 497

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 498 CACATCCAGCGCATGACACCGGACATTAAGCGGACATTAAAGAGCATATGTTCTTCAGG 557
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 558 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCGCATATGAGAGTAT 617
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTyr 200
DB 618 AACCCGAGGTGGGCTACTGCAGGGAGCTGAGCCACATCGCCGCTGTCTCTCTAT 677
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 678 CTTCCTGAGGAGATGCATTCTGGGCACTGTGCACTGCTGGCCAGTAGAGGCACTCC 737
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 738 CTGCAGGATTTCAAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAGAG 797
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 798 CATGTGTAGCCACGTCAACAATCCAAAGACATGGGGCATCAGAGCAAGAAAGATCTATGT 857
QY 261 GlyLysCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGly----- 276
DB 858 GGGCAGTGTTCCTTGGCTTAGGCTGCTCCTCATCCGGATATTGATTGACGGGGTAAGAGGCAT 917
QY 276 ----- 276
DB 918 AGGAGACCTCGCTCAGGGAACCTTCTGCTGCTGCAAGTGCCTGCTTCCCAAGCCCGGG 977
QY 276 ----- 276
DB 978 GGTCTGGCTCACTCCAGCCCAAGAGGCTCAGCGGGTCCCAAGACACACAAGCA 1037
QY 276 ----- 276
DB 1038 AAACCTCTGCCCAGAGGGGTCAATCCCAAGGCAATGGCTGGGGCTCAGGCCAGCTCA 1097
QY 276 ----- 276
DB 1098 TGGGCAAGCTGGGCCAGGACCCGACTTGAGAGGGCTCAGGGAAGCCTCAAGCCCTGGGCA 1157
QY 276 ----- 276
DB 1218 AGACCTGTGTGACCAGCGCTCTGAGGGCTCAGCGACCTCATGGGGAAGTCACTG 1277
QY 276 ----- 276
DB 1278 ACTCTGAGACTGAAGCCCGAGCGTGGCAGCTCGAGCCACCAAGCCCAAGCTGGAAGGA 1337
QY 277 -----IleSerLeuGlyLeuThrLeuArgLeuTr 286
DB 1338 CCAGGTTCTTTCACACTGCTGTCCCAAGATCTCTCGGGCTCAGCCCTGCGCTGTG 1397
QY 286 PAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLys 306
DB 1398 GGAAGTGTATCTGTGAGAAAGCGCAAGCGCTTGATGCCGATTAACAAGATCGCCTTAA 1457
QY 306 sValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAs 326
DB 1458 GGTTCAGCAGAGCGCTCAGAAAGACGTCCAAGTGTGGCCCGTGGGCAAGTTTTCGA 1517
QY 326 nArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSe 346
DB 1518 CCGGTTCGTTGATACCTGGGCGAGGATGAGACACTGTGCTCAAGCATCTTAAGGCGCTC 1577
QY 346 mMetLysLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGlnG 366

Db 1578 TATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCGACCAACCCGAGCAAG 1637
QY 366 YSERSERAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAs 386
Db 1638 GTGTCGGCATCCAGGCTGTGCGGCTTCACTGGCGGGAAGACCCTCTGCAAGGGGA 1697
QY 386 PARGLAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSerPr 406
Db 1698 CAGCAGGCCCCCTCCAGGCCCAAGCCCGGTTCCCGCGGCATTGGTCAGCTTCCCT 1757
QY 406 OProArgAlaProArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTy 426
Db 1758 GCCACGGGACCTCGTTCTTCCACACCCCTGTCTGTGGGCTGTCCGGGAAGACACCTA 1817
QY 426 rProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySe 446
Db 1818 CCTGTGGGCACCTCAGGGGTGTGCCAGCCCGGCTGTGCTCAGGAGGACCTCAGGGTTT 1877
QY 446 rTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnG 466
Db 1878 CTGGAGATTCTGCAAGTGAACCTTCATGCCCCCGCTCCCAACGGACCTGACGTAGAAGG 1937
QY 466 YProTrpPheArgHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnG 486
Db 1938 CCCTTGTTCGCCCATATGATTTACAGACAGAGCTGTGGGTCCGTGCCATATCCACAGA 1997
QY 486 uAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPh 506
Db 1998 GGACCACTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAGCGGGTGAGATCGGCTTT 2057
QY 506 eAlaAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGluGlnPr 526
Db 2058 CACTGCACCCAGCACTGATTCGACAGGGCACCCCCTTCAGAGCTTAGGAGCAAGCAACA 2117
QY 526 oCyAlaProThrSerGlyProCysLeuGlyLeuHisLeuGlnSerSerGlnPhePr 546
Db 2118 GTGTGCTCCCACTCAGGGCTGTGCTGTGGGCTCCACTTGGAAGATTCTCAGTTCCC 2177
QY 546 oProGlyPhe 549
Db 2178 TCCAGGCTTC 2187

RESULT 11
AX775951 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775951
DEFINITION Sequence 221 from Patent WO03048202.
ACCESSION AX775951
VERSION AX775951.1 GI:32693669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Matsuda, A. and Muramatsu, S.
AUTHORS NF-KB activating gene
TITLE Patent: WO 03048202-A 221 12-JUN-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source 1. .2647
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11675.1"
/db_xref="GI:32693670"
/translation="MDVVEVAGSWMAQEREDIMKYEKGHRAGLPEDKGPKEPSYNN
NVNHLGI VHETELPELTAREAKQIRREISRSKSWDMLDWEKTKSSRLIDRAYKGM
PMNIRGPMWSVLNTEEMKLNKNGRYQIMKEKGRSSEHQRIDRDVSGTLRKHIFER
DRYGTQRRLHILLAYEEYNDEVGYCRDLSHIALFLYLPEBDAFWALVQLASER

HSLOGFHS PNGSTVQGLDQDEHVATSQPKTMGHQISLGLTLRLMDVYLVEGEALM
PI TRIAFKYOQKRLTTKTSRCGPWAFPCNRPVDTWARDEDVYLKILRASMKLTRLKQGD
LPPPAKPEOGSSASRPVPASRGKTLCKGDRQAPGPAPAPRPPIWSASPPAPRSST
PCPGAVREDDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDLVGVGPMFRHY
DFQSCWTRAI SQEDQLAPCWQAEHPARVRSAFAAPSTDSDQGTPEFRARDEQPCAPT
SGPCLGHLHLESQFPQPF"

ORIGIN
Alignment Scores:
Pred. No.: 5, 3e-98
Score: 2835.00
Percent Similarity: 95.81%
Best Local Similarity: 95.81%
Query Match: 95.04%
DB: 6
Gaps: 1
Length: 2647
Matches: 526
Conservative: 0
Mismatch: 1
Indels: 22

US-10-071-838-2 (1-549) x AX775951 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 755 ATGGACGTGTGAGAGTCCGCGGCAAGTTGGTGGCACAAGACCGAGACATCATATG 814
QY 21 LysTyrGlyLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 815 AAATACGAAGAAGGACACCGAGCTGGCTCCAGAGACAGAGGGCCCTTAAGCTTTTCCA 874
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 875 AGCTACACAACAACGTCATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 935 ACTGGCGGGAGGCGGAAGCAAAATTCGCGGGAGATCAAGCAAGCAAGTGGTGGAT 994
QY 81 MetLeuGlyAspTrpGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAGAAAGCTATGATCGAGCTTACAAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1055 GGAATGCCCATGAACATCCGGGGCCCGATGTGCTCAGTCTCTGAAACATGAGGAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGCTCATCTAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATTCGACCGGAGCTGAAGCGGACATTAAAGAAAGCATATATTCTCAGG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 1235 GATCGATACGGAACCAAGCAGCGGGAACCTACCTCACATCTCTGGCATATAGAGATAT 1294
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 1295 AACCCGGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCGCTTGTCTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCCTGAGGAGATGATCTTGGGCACTGTGCAGCTGTGGCCAGTGAAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1415 CTGCAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 1475 CATGTGTAGCACGTACAAACCAAGACCATGGGCATCAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1517 -----ATCTCTCTCGGG 1528

Qy 281 LeuThrLeuAArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300
Db 1529 CTCACCTTGGCCCTGTGGACGTGTATCTGGTAGAAGCGCAACAGCGCTTGATGCCGATA 1588
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1589 ACAAGAAATCGCCTTTAAGGTTTCAGCAGAAAGCGCTTCAGAAAGACGTCCAGGTGTGCCCG 1648
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1649 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1708
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCCA 1768
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1769 GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCGCTGTGCCGCTTCACGTGGCGGGAAG 1828
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGCAAGGGGAGACAGGCAAGCCCTCCAGGCCCAACAGCCCGGTTCCCGCGGCC 1888
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1889 ATTGGTCAGCTTCCCGCCACGGGCACCTCGTTCTTCCACACCCCTGTCTGTGGGGCT 1948
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGTGCTCAG 2008
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 2009 GAGGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCTCATGCCCGCTCCCAACG 2068
Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 2069 GACCTGACGTAGAGGGCCCTTGCTTCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 2128
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 2129 CGTGCCATATCCAGAGAGCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 2188
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 2189 CGGGTGAGATCGGCTTTCGCTGCACCCAGACACTGATTCGAGCAGGACACCCCTTCAGA 2248
Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 2249 GCTAGGAGCAGAACAGCCGTGTCTCCACCTCAGGCGCTTGCTCTGCCGCTCCACTTG 2308
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335
RESULT 12
AX775949 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775949
DEFINITION Sequence 219 from Patent WO03048202.
ACCESSION AX775949
VERSION AX775949.1 GI:32693667
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Matsuda, A. and Muramatsu, S.
AUTHORS NF-kB activating gene
TITLE Patent: WO 03048202-A 219 12-JUN-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
Source Location/Qualifiers
1..2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11674.1"
/db_xref="GI:32693668"
/translation="MDVEVAVAGSMVAQEREDIMKYEKHRAGLPBDKPKPFRSYNN
NVDLGIIVHETPLPTAREAKQIRREISRSKWVMDLGDMEKYSSRKLIDRAYKGM
PMNIRGPMWSVLNTEBEMKLNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFR
DRYGTKORBLHLILAYEYNPEVYCRDLSHIALFLYLPEEDAFWALVOLLASER
HSLQGFHSFNGSTVQGLQDOEHVVA TSQPTMGHQLSLGLTLRLMDVYLVEGEQALM
PITRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDEDTVLKHLRASMKLTRKGD
LQPPAKPEQSSASRPVPASRGKTLCKGDRQAPPGPPAPRPPIWSASPRAFRST
PCPGAVREDTYPVGTQVPSPALAGGPGQSMRFLQWNSMPRLPTDL DVEGPFWRHY
DFRQSCWVRALISQEDQLAPCWQAEHPAERVRSAFAAPSTSDQGPFRARDEQCAPT
SGPCLGLHLESSQFPFPGF"
ORIGIN
Alignment Scores:
Pred. No.: 1.06e-97 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX775949 (1-2647)
Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 755 ATGACGTGTAGAGGTCCGGGAGTTGTGGGCCAACAGCGAGACATCATTAATG 814
Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 815 AAATAAGAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCCTTAAGCCTTTTGA 874
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 875 AGCTAACAAACAACACGTGCATCTTGGGAGATTGATGAGACGAGCTGCTCTCTG 934
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 935 ACTGCCGGGAGGCGCAACAATTCCGCGGGAGATCAGCCGAAAGACAAAGTGGTGAT 994
Qy 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 995 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 1054
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db 1055 GGAATGCCCATGAACATCCGGGGCCGATGTGTCAGTCTCTGAACTGAGGAATG 1114
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGCAAGGTCTCTGAG 1174
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGAAAGCATATCTTCAGG 1234
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCTGCGCATATGAGAGATAT 1294
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 1295 AACCCGAGGTGGGCTACTGACGAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 1354
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

Db	1355	CTTCCTGAGGAGGATGATTCTGGGCACGTGTGCAGCTGTGGCCAGTAGAGGACACTCC	1414
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGly	240
Db	1415	CTGCAGGGATTTCACAGCCCAATGGCGGGACGTCCAGGGGCTCCAAAGACCACAGAGAG	1474
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	1475	CATGTGTAGCCACGTCAACAACCAAGACCATGGGCATCAG-----	1516
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	1517	-----ATCTCTCTCGGG	1528
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
Db	1529	CTCACCCCTGCGCTGTGTGGAGCGTATCTGGTAGAAGGCGAACAAGCGTTGATGCCGATA	1588
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	1589	ACAAGAATCGCCTTTAAGGTTTACAGACAGAGCGCTCACAGAGAAGCTCCAGGTGGCCCG	1648
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1649	TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC	1708
QY	341	LysHisLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1709	AAGCATCTTAGGGCTCTATGAGAAACTAACAGAAAGCAGGGGAGCCTGCACACCCCA	1768
QY	361	AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
Db	1769	GCCAAACCCGACGAAGGGTCTGTGGCATCCAGGCCCTGTGCGCTTCACGTGGCGGGAG	1828
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1829	ACCCTGTGCAAGGGGACAGGACGGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC	1888
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla	420
Db	1889	ATTGGTCACTTCCCGGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGGTGGGCT	1948
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1949	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGCCCTGCTCAG	2008
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	2009	GGAGGACCTCAGGGTTCCTGGAGATTCTCGACGTGGAATCCATGCCCGCCCTCCCAACG	2068
QY	461	AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTyrVal	480
Db	2069	GACCTGGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGACTGCTGGGTC	2128
QY	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAlaGlu	500
Db	2129	CGTGCCATATCCAGAGGAGACCAAGTCGGCCCTGCTGGCAGGGCTGAACACCTGCGGAG	2188
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
Db	2189	CGGGTGAGATCGGCTTTCCTGTCACCCAGCACTGATTCGACCAAGGGCAACCCCTTCAGA	2248
QY	521	AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
Db	2249	GCTAGGAGACGAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTGCGGCTCCACTTG	2308
QY	541	GluSerSerGlnPheProProGlyPhe	549
Db	2309	GAAAGTTCTCAGTTCCTCCAGGCTTC	2335

RESULT 13
AX833552

LOCUS	AX833552	2647 bp	DNA	linear	PAT 15-DEC-2003
DEFINITION	Sequence 676 from Patent EP1347046.				
ACCESSION	AX833552				
VERSION	AX833552.1	GI:39919687			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.				
TITLE	Full-length cDNA sequences				
JOURNAL	Patent: EP 1347046-A 676 24-SEP-2003; Research Association for Biotechnology (JP)				
FEATURES	location/Qualifiers				
source	1..2647				
ORIGIN	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
Alignment Scores:					
Pred. No.:	1.06e-97	Length:	2647		
Score:	2827.00	Matches:	525		
Percent Similarity:	95.63%	Conservative:	0		
Best Local Similarity:	95.63%	Mismatches:	2		
Query Match:	94.77%	Indels:	22		
DB:	6	Gaps:	1		
US-10-071-838-2 (1-549) x AX833552 (1-2647)					
QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20		
Db	755	ATGACGCTGTAGAGGTCCGGCACTTGTGTGGGCAAGAAGCAGAGACATCATATATG	814		
QY	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40		
Db	815	AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGCCCTAAGCCTTTTGA	874		
QY	41	SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGlnLeuProProLeu	60		
Db	875	AGCTACAACAACAACGTGCATCTTGGGATTTGTATGATGAGACGAGAGCTGCTCTCTG	934		
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80		
Db	935	ACTGCGGGGAGCGCAAGCAATTCCGGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT	994		
QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100		
Db	995	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCAAG	1054		
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnIleMet	120		
Db	1055	GGAATGCCATGAACATCCGGGGCCGATGTGTCAGTCTCTCTGAACAACGTAGAAATG	1114		
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140		
Db	1115	AAATTGAAAAAACC CGGAAGATACCAATCATGAAGAGAAAGGCAAGAGTCATCTGAG	1174		
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160		
Db	1175	CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAAGAAACATATATTCTTCAGG	1234		
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180		
Db	1235	GATCGATTACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGCGCATATGAGAGAT	1294		
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200		
Db	1295	AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCGCCTTGTCTCTCTAT	1354		

```
QY      201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
      |||
Db      1355 CTTCCGTGAGGAGATGCAATTCTGGGCACCTGGTGCAGCTGCTGCCAGTGAAGGCACTCC 1414
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
      |||
Db      1415 CTGCAGGATTTCACAGCCCAATGGCGGAGCCGTCACAGGGGCTCCAGAACCAACAGAGAG 1474
QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
      |||
Db      1475 CATGTGTAGCCACGTCACAACCAAGACCATGGGGCATCAG----- 1516
QY      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
      |||
Db      1517 -----ATCTCTCTCGGG 1528
QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
      |||
Db      1529 CTCACCCCTGCCCTGTGGGACGTGTATCTGTAGAGAGCGCAACAGCGCTTGATCCGATA 1588
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
      |||
Db      1589 ACAAGAATCGCCTTTAAGTTTACAGACAGAGCGCTCACGAAGACGTCACAGTGTGGCCG 1648
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
      |||
Db      1649 TGGGCACGTTTGTGCAACCGTTCGTTGATACCTGGGCCACAGGATGAGACACTGTGCTC 1708
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
      |||
Db      1709 AAGCATCTTAGGGCCTTATGAAGAACTAACAGAAAGAGGGGACCTGCAACCCCA 1768
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
      |||
Db      1769 GCCAAACCCGAGCAAGGCTCTCGGCATCCAGGCTGTGCGCGCTTCACGTGGCGGAG 1828
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
      |||
Db      1829 ACCCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGCCCAACAGCCCGCTTCCGCGGCC 1888
QY      401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
      |||
Db      1889 ATTTGTCAGCTTCCCGCCACGCGCACCTCGTCTTCCACACCTGCTGTGGGGCT 1948
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
      |||
Db      1949 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGTGCCACGCCGCCCTGCTCAG 2008
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
      |||
Db      2009 GGAGGACCTCAGGGTTCCTGAGATTCCTGCAGTGAATCCATGCCCGCTCCCAACG 2068
QY      461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
      |||
Db      2069 GACCTGACGTAGAGGGCCCTGTGTTCCGCAATTATGATTTCAGACAGAGCTGCTGGGTC 2128
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
      |||
Db      2129 CGTGCCATATCCAGAGAGACCAAGCTGAGCCCTGTGTCAGAGGCTGAACACCCCTGCGAG 2188
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
      |||
Db      2189 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGCAACCCCTTCA 2248
QY      521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
      |||
Db      2249 GCTAGGAGACGAACAGCGGTGTGCTCCCACTCAGGGGCTTGCTGCGGCTCCACTTG 2308
QY      541 GluSerSerGlnPheProProGlyPhe 549
      |||
Db      2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335
```

RESULT 14

AK095385

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK095385 2647 bp mRNA linear PRI 30-JAN-2004
Homo sapiens cDNA FLJ38066 fis, clone CTONG2015330, moderately
similar to H.sapiens mRNA for tre oncogene.

AK095385
AK095385.1 GI:21754631
oligo capping; fis (full insert sequence).
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Oca,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hokuwa,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Ohtani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

JOURNAL

PUBMED

REFERENCE

AUTHORS

Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2647)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 2647

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2015330"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="Cloning vector: pME18SFL3"

Alignment Scores:

Pred. No.:
Score: 1.06e-97
Percent Similarity: 2827.00
Best Local Similarity: 95.63%
Query Match: 95.63%
DB: 94.77%

Length: 2647
Matches: 525
Conservative: 0
Mismatch: 2
Indels: 22
Gaps: 1

US-10-071-838-2 (1-549) x AK095385 (1-2647)

QY	1	MetAspValValGluValAlaGlySerTTPTrpAlaGlnGluArgGluAspIleIleMet	20
DB	755	ATGACCTGGTAGAGTCCGGCGGACGTTGGTGGCACAAAGAGCGAGGACATCATATG	814
QY	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	815	AAATACGAAAGGACACCGAGCTGGCTCCAGAGAGCAAGGGGCTTAAGCCTTTTGA	874
QY	41	SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu	60
DB	875	AGCTACACACAAAGCTCGATCATTTGGGATTGTACATGAGACGAGCTCCTCTG	934
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
DB	935	ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT	994
QY	81	MetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
DB	995	ATGCTGGGAGACTGGGAGAATACAAAGCAGCAGAAAGCTCATAGATGAGCGTACAG	1054
QY	101	GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGluMet	120
DB	1055	GGAATGCCCATGAAACATCCGGGGCCGATGTGGTCTGCTCTGAACTGAGGAAATG	1114
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
DB	1115	AAGTTGAAAAACCCCGGAAGATTACCATCATGAAGAGAGGCAAGAGTCAATGAG	1174
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
DB	1175	CACATCCAGCGCATGACCGGAGCTTAAGCGGACATTAAAGGACATATATCTTCAGG	1234
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
DB	1235	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT	1294
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr	200
DB	1295	AACCCGAGGTGGCTACTGACGGACCTTGAGCCACATCGCCGCTTCTCTCTAT	1354
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
DB	1355	CTTCTGAGGAGATGATCTTGGGCACTGTGCAAGCTGTGGCCAGTGAGGCACTCC	1414
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
DB	1415	CTGACAGGATTTCAACAGCCCAATGCGGACCTGCCAGGGCTCCAAAGCAACAGGAG	1474
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys	260
DB	1475	CATGTGTAGCCACGTCAACCAAGCAATGCGGCACTCAG-----	1516
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280

DB	1517	-----ATCTCTCTCGGG	1528
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle	300
DB	1529	CTACCTCGCGCTGTGGGACGTGTATCTGTGTGAAGCGCAACAGCGGTGATGCCGATA	1588
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
DB	1589	ACAAGATCCCTTTAAAGTTCAAGCAGAGAGCCCTCACGAAAGACGTCCAGGTGTGCCG	1648
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu	340
DB	1649	TGGGCACTTTTTCACACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTCTC	1708
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
DB	1709	AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCACACCCCA	1768
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
DB	1769	GCCAAACCCGACCAAGGTCGTCCGATCCAGGCTGTGCCGCTTACCTGTGGGAG	1828
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
DB	1829	ACCCTTGCAGAGGGGACAGGAGGCCCTCCAGGCCACCAAGCCCGCTTCCCGCGCC	1888
QY	401	IleTyrSerAlaSerProProArgAlaProArgSerThrProCysProGlyAla	420
DB	1889	ATTGCTCAGCTTCCCGCCAGCGGACCTGTTCTTCACACCTGTCTGTTGGGCT	1948
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
DB	1949	GTCCGGGAAGACACTACCTGTGGGACACTCAGGTTGTCACGCCCGCTGGCTCAG	2008
QY	441	GlyGlyProGlnGlySerTyrPArgPheLeuGlnTyrAsnSerMetProArgLeuProThr	460
DB	2009	GGAGGACCTCAGGTTCTCTGAGATTCTCGAGTGAATCCATGCCCCCGCTCCCAAG	2068
QY	461	AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
DB	2069	GACCTGACGTAGAGGGCCCTGTGTTCCGCCATATGATTGACACAGAGCTGGGCTC	2128
QY	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAlaGlu	500
DB	2129	CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGACACCTGCGGAG	2188
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
DB	2189	CGGGTGAATCGGCTTTCGTGACCAACCACTGATTCGACCAAGGCAACCCCTTCAGA	2248
QY	521	AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
DB	2249	GCTAGGAGCAGAACCGGTGTCTCCACCTCAGGGCTTGCCTGTGGGCTTCACTTG	2308
QY	541	GluSerSerGlnPheProProGlyPhe	549
DB	2309	GAAAGTTCTCAGTTCCCTCAGGCTTC	2335

RESULT 15
BC058890
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

2111 bp mRNA linear PRI 08-OCT-2003
Homo sapiens cdna clone MGC:64921 IMAGE:5744726, complete cds.
BC058890.1 GI:37590617
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Klausner, R.D., Collins, F.S., Wagner, L., Schenker, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedlin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skaleka,U., Smalhus,D.E.,
Schererch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2111)
Strausberg,R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zyderdun, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov
Series: IRAK Plate: 119 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984.
location/Qualifiers
1. .2111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:64921 IMAGE:5744726"
/cisue_type="Brain, adult medulla"
/clone_id="NIH MGC_119"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
85. .1125
/codon_start=1
/product="Unknown (protein for MGC:64921)"
/protein_id="AAH58890.1"
/db_xref="GI:37590618"
/translation="MDVVEVVGSMVAQEREDIMKYEKHRAGLPEDKPKFSGSYNN
NIDHFGMOHETELPPLTAREVKQIRREISRKSKWVKMLGEMPTYNKSRKIDRAQGI
PMNIRGPMWSVLNIEIKNPNRGYQIMKEKRSSEHIQMDLDVSGTLRRHIFFR
DRYGTQKRELYTLLAYEBYKNPEVGYCRDLSHIALFLYBEPDAFWALVOLLASER
HSLQGFHSPNGSTVQGLQDQOEHVATSLPKTMHQDKDLCCGCSLGLIRILIDG
ISLGLTLRLMDVYLVGEQAFMLITKSPLRFSRSASRRRPVARGHVFGTSLMPGPG
MTLCSILGPI."
misc_feature
379. .1011

/notes="TBC; Region: Domain in Tre-2, BUB2p, and Cdc16p.
probable Rab-GAPs. Widespread domain present in Gyp6 and
GYP7, thereby giving rise to the notion that it performs a
GTP-activator activity on Rab-like GTPases"
/db_xref="CDD:smart00164"
ORIGIN
Alignment Scores:
Pred. No.: 3.05e-94 Length: 2111
Score: 2732.00 Matches: 513
Percent Similarity: 93.44% Conservative: 14
Best local Similarity: 90.96% Mismatches: 22
Query Match: 91.59% Indels: 16
DB: Gaps: 1
US-10-071-838-2 (1-549) x BC058890 (1-2111)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 85 ATGGAAGCTGTAGAGTGTGCTGGTGTGTTGGTGGGACCAAGAGCGAGACATCATTTATG 144
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 145 AATATCAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCCTTTTGA 204
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 205 AGCTACAAACAACATCATCTGATCACTTGGGATGACATGAGACGAGCTGCTCTCTG 264
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 265 ACTGCGCGGAGGTGAAGCAATTCGGCGGAGATCAGCCGAAGCAAGTGGTGA 324
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 325 ATGCTGGAGATGGACACCTACAAAACACAGAAAGCTCATGATCGAGCTACCAAG 384
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 385 GGAATTCATGAAATCCGCGGCGGCGGATGTGTCACTCTCTTGAACATTGAGGAATC 444
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 445 AAGTTGAAAACCCCGAAGATACCATCATGAAAGAGAAAGGCAAGAGTCACTTGA 504
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 505 CACATCCAGCAGATGACCTGAGCAGTAAAGTGAATTAAGAGCAGATATCTTCAG 564
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLys 180
DB 565 GATCGATACGGAACCAAGCAGCGGAGACTTATCATCTCTCTGCGGTATGAGAGTAT 624
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
DB 625 AACCCGAGAGTGGCTACTGACGAGACTGAGCCACATCGCCCTTGTCTCTTAT 684
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 685 CTTCCTGAGAGATGATCTTGTGGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 744
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 745 CTGACAGGATTTCAAGCCCAATGGCGGACCTGCCAGGGGCTCAAGAACCAAGAGAG 804
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 805 CATGTGTAGCCAGCTCACTACCAAGACATGTGCGATCAGAGCAAGAAAGATCTATGT 864
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
DB 865 GGGCAGTGTTCCTTAAAGCTGCTCATCCGATATTTGATGAGGATCTCTCTCGGG 924

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGluAlaLeuMetProIle 300
 Db 925 CTCACCTCGCGCTCTGGGACGTGTATCTGTAGAAAGCGACAGCGCTTCATGCTGATA 984
 QY 301 ThrArgIleAlaPheLysValGluGluLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 985 ACAAA-ATCGCTTTAAGGTTCAAGCAGAAAGCGCTCACGAGACGTCAGGCTGTGCCG 1043
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1044 TGGGACGTTTGTGAACCGGTCCTGTGATGCTGTGGCCAGGATGATGACACTGTGCTC 1103
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGluGlyAspLeuProProPro 360
 Db 1104 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1163
 QY 361 AlaLysProGluGluGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 1164 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCCAGCTTCCAGGGCGGGAAG 1223
 QY 381 ThrLeuCysLysGlyAspArgGluAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1224 ACCCTCTGCAAGGGGAG 1283
 QY 401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
 Db 1284 ATTTGGTCAGCTTCCCGCCAGCGGCACTGTTCTTCCACACCTGCTGTGGGGCT 1343
 QY 421 ValArgGluAspThrTyrProValGlyThrGluGlyValProSerProAlaLeuAlaGln 440
 Db 1344 GTCCGGAGAGACACTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGGCTCAG 1403
 QY 441 GlyGlyProGluGlySerTyrPheArgPheLeuGlnTyrAsnSerMetProArgLeuProThr 460
 Db 1404 GGAGGACCTCAGGGTCTCTGAGATCTCGCAGTGAATCCATGACCCCGCTCCCAACG 1463
 QY 461 AspleuAspValGluGlyProTyrPheArgHisTyrAspPheArgLysSerCysTrpVal 480
 Db 1464 GACCTGACGTAGGGACCTTGTGTCGCGCTTATGATTTCAAGACAGAGCTGTGGGTC 1523
 QY 481 ArgAlaIleSerGluGluAspGluLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
 Db 1524 CGTGCCATATCCAGAGAGACAGCCGCGCACCTGCTGGCAGGCTGMAACACCTGCGGAG 1583
 QY 501 ArgValArgSerAlaPhe----- 506
 Db 1584 CGGGTGAGATCGGCTTCACTGCACCTGAGCCACAACGTGGGCATGACTTCCCGGCCCTG 1643
 QY 507 -AlaAlaProSerThrAspSerAspGluGlyThrProPheArgAlaArgAspGluGlnPr 526
 Db 1644 CAGTGGCGCCAGCAGCTGATTCGACCAAGGCGACCCCTTCAGAGCTAGGACGAAACAGCA 1703
 QY 526 oCybAlaProThrSerGlyProCysLeuCysGlyLeuHisLeuGluSerSerGlnPhePr 546
 Db 1704 GTGTGCTCCCACTCAGGGCTTGTGCTTGTGGGCTTCCACTTGAAAGTTCTCAGTTCCC 1763
 QY 546 oProGlyPhe 549
 Db 1764 TCCAGGCTTC 1773

Search completed: February 4, 2005, 10:14:09
 Job time : 6882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 05:10:43 ; Search time 753 Seconds
(without alignments)
3827.267 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVVEVAGSWAQEREDIIIM.....TSGPCLGHLHSSQFPFPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US10071838/runat_03022005_071217_179/app_query.fasta_1.711
-DB=N_Geneseq_23Sep04 -QFMT=fastcap -SUFFIX=mg -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071838@cgn_1_1_470 @runat_03022005_071217_179 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- N_Geneseq_23Sep04:*
- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	6	ABSS54706
2	2983	100.0	1964	10	ADC37382
3	2978	99.8	1993	3	AAAI5001
4	2975	99.7	1964	10	ADC37384
5	2963	99.3	2072	5	ABX71403
6	2907.5	97.5	2146	6	ABSS54707

7	2906.5	97.4	2304	10	ADC37380	Adc37380 Nuclear f
8	2898.5	97.2	2304	10	ADC37378	Adc37378 Nuclear f
9	2850	95.5	2084	6	ABZ11810	Abz11810 Human pol
10	2850	95.5	2084	12	ADM44328	Adm44328 Novel hum
11	2835	95.0	2647	10	ADC37388	Adc37388 Nuclear f
12	2827	94.8	2072	6	ABZ11809	Abz11809 Human pol
13	2827	94.8	2072	12	ADM44327	Adm44327 Novel hum
14	2827	94.8	2647	10	ADC37386	Adc37386 Nuclear f
15	2827	94.8	2647	11	ADM01991	Adm01991 Human cdn
16	2784	93.3	1752	6	ABST78751	Abst78751 DNA encod
17	2773	93.0	1862	6	ABSS54708	Abss54708 CDNA enco
18	2191.5	73.5	7856	5	AA85531	Aa85531 DNA encod
19	2191.5	73.5	7856	5	AA888213	Aa888213 DNA encod
20	2191.5	73.5	7878	6	ABL64684	Ab164684 Stomach c
21	2135.5	71.6	8435	4	AA159633	Aa159633 Human pol
22	2135.5	71.6	8435	4	AA159634	Aa159634 Human pol
23	2135.5	71.6	8435	10	ADC31883	Adc31883 Human nov
24	2133.5	71.5	8180	5	AA888212	Aa888212 DNA encod
25	2133.5	71.5	8201	2	AAT12170	Aat12170 pJG4-5-CD
26	2133.5	71.5	8201	10	ADE85053	Ade85053 Farnesyl
27	2133.5	71.5	8284	5	AA85535	Aa85535 DNA encod
28	2133.5	71.5	8408	4	AA157848	Aa157848 Human pol
29	2133.5	71.5	8420	4	AA157847	Aa157847 Human pol
30	1795.5	60.2	3396	5	AA885534	Aa885534 DNA encod
31	1795.5	60.2	3904	5	AA888214	Aa888214 DNA encod
32	1732	58.1	2838	5	AA887029	Aa887029 DNA encod
33	1732	58.1	2838	5	AA884009	Aa884009 DNA encod
34	1732	58.1	2838	5	AA887522	Aa887522 DNA encod
35	1533.5	51.4	2424	5	AA887523	Aa887523 DNA encod
36	1507.5	50.5	9805	4	AA836324	Aa836324 Human car
37	1507.5	50.5	9805	4	AA105169	Aa105169 Human rep
38	1507.5	50.5	9805	4	ABL98056	Ab198056 Human tes
39	1507.5	50.5	9805	10	ADE47018	Ade47018 Human car
40	1500.5	50.3	9404	4	AA836326	Aa836326 Human car
41	1500.5	50.3	9404	4	AA105171	Aa105171 Human rep
42	1500.5	50.3	9404	4	ABL98058	Ab198058 Human tes
43	1500.5	50.3	9404	10	ADE47020	Ade47020 Human car
44	1469	49.2	9804	4	AA836329	Aa836329 Human car
45	1469	49.2	9804	4	AA105173	Aa105173 Human rep

ALIGNMENTS

RESULT 1					
ID	ABSS54706	standard;	CDNA;	1964	BP.
XX	AC	ABSS54706;			
XX	DT	28-NOV-2002	(first entry)		
XX	DE	CDNA encoding human PRC17 protein.			
XX	KW	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;			
XX	KW	breast cancer; cytostatic; chromosome 17q11-12; gene; ss.			
XX	OS	Homo sapiens.			
XX	FH	Key	Location/Qualifiers		
FT	CDS		1..1650		
FT			/*tag= a		
FT			/product= "Human PRC17 protein"		
XX	PN	WO200262958-A2.			
XX	PD	15-AUG-2002.			
XX	PF	08-FEB-2002; 2002WO-US003457.			
XX	PR	08-FEB-2001; 2001US-0267615P.			
XX	PA	(TULA-) TULARIK INC.			
XX					

PI Li J, Powers S, Xiang P, Peng Y;
XX
DR WPI; 2002-706902/76.
XX P-PSDB; ABG70736.

PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PS or expression relative to normal, for example cancer.
XX Claim 24; Page 62; 78pp; English.

CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence represents the human PRC17 gene
CC located on chromosome 17q11-12. This sequence encodes the human PRC17
XX protein of the invention

SQ Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.42e-128

Score: 2983.00 Length: 1964

Percent Similarity: 100.00% Matches: 549

Best Local Similarity: 100.00% Conservative: 0

Query Match: 100.00% Mismatches: 0

DB: 6 Indels: 0

US-10-071-838-2 (1-549) x ABS54706 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGACGTCGTAGAGGTCCGGCGCAGTGTGTGGGCACAAAGACGAGACATCATATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAGGACACCGAGCTGGCTGCAGAGACAAAGGGCCCTTAAACCTTTTCA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 121 AGCTACACACACACCTGCATCATTTGGGATGTATGATGAGACGGAGCTGCTCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTGCGCGGAGGCGAAGCAATTCCGGGAGATCAGCCGAAGAAGCAAGTGGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTGGAGAATAACAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCCATGTGTCACTCTCTGACATTTAGGAAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAAACCCCGAAGATACCATGATGAAGAGAGGAGGAGATCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATGACCGGAGCTAAGCGGACATTAAAGCAATATATCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 481 GATCGATACGGAACCAAGACGCGGACTACTCCACATCTCTGCGATATGAGAGATAT 540
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200

Db 541 AACCCGAGGTGGGCTACTGCAGGGACCTGAGCCACATGCCGCCCTTGTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTCTCTGAGGAGATGATCTTGGGCACTGTGTGACGTGCTGGCCAGTAGAGGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 661 CTGCAGGATTTCAACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 721 CATGTGGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGACAAAGAAATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGGAATTGATTGACGGGATCTCTCTCGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 841 CTCACCTGGCGCTGTGGACGTATCTGTGAAGAGCGAAGCGGCTGATGCCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAAGAAATCGCCTTTAAGTTACAGACAAAGCCCTCACAAAGACGTCCAGTGTGCCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 961 TGGGCACTTTTGCACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1021 AAGCATCTTAGGCGCTCTATGAGAAACTAACAAAGAAAGCAGGGGACCTGCCACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1081 GCCAAACCCGACAAAGGTCGTGGCATCCAGGCCCTGTGCGGCTTCACTGCGGGAAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1141 ACCCTTCGCAAGGGGACAGGACAGGCCCTCCAGGCCCAAGCCCGGTTCCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
Db 1201 ATTGTGTAGCTTCCCGCCCAAGGCACTCGTTCTTCCACACCTGTCTGTGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1321 GAGGACCTCAGGGTTCCTGAGATTCTTCAGTGAAGCACTCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1381 GACCTGACGTAGAGGGCCCTTGTTCGCCCATTTATGATTTCAGACAGAGCTGGGTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1441 CGTGCATATCCCAAGAGACCAAGCACTGGGCCCTGTGGCAGGCTGAACACCTGGCAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1501 CCGGTGAGATGGCTTTCGTGACCCAGACATGATTCCAGCCAGGAGCAACCCCTTCAAG 1560
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1561 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCCTTGCCCTTCCCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1621 GAAAGTTCTCAGTTCCTCCAGGCTTC 1647

RESULT 2
ADC37382
ID ADC37382 standard; DNA; 1964 BP.
XX
AC ADC37382;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
XX WPI; 2003-505282/47.
DR P-PSDB; ADC37383.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischaemic disorders.
XX
XX
PS Claim 4; SEQ ID NO 215; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
XX
SQ Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.42e-128 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-071-838-2 (1-549) x ADC37382 (1-1964)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGACGCTGTAGAGGTGCGGGCAGTGTGGGCAACAAGAGCGAGACATCATATATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAGGAGACACCGAGCTGGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTGA 120
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60

Db 121 AGCTACAACAACAGCTGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAGCAAGTGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTGGGAGAATAACAAAGCAGAGAAGCTCATGATCGAGCGTACAAG 300
QY 101 GlyMetProMetAsnIleArgGlyPrometTrpSerValLeuLeuAsnIleGluIleMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCAAGTCTCTCTGAACATTGAGGAAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAACCCCGAAGATACAGATCATGAAGAGAGGCAAGGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGAAAGCATATATTCTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLys 180
Db 481 GATCGATACGGAACCAAGCAGCGGAGACTACTCCACATCTCTCGCATATGAGAGATAT 540
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 541 AATCCGGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCCTGAGGAGATGATCTTGGGCACTGTGTGACGCTGTGGCCAGTGAAGGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 661 CTGACGAGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCAACAGAGAG 720
QY 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysLysAspLeuCys 260
Db 721 CATGTGTAGCCAGTCAACAACCAAGACATGGGGCATCAGAGCAAGAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGAGATATGATGACGGATCTCTCGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlyGlyGlnAlaLeuMetProIle 300
Db 841 CTCACCTGCGCTGTGGAGCGTGTATCTGTGAAGGCGAAGCGGCTGTATGCCGATTA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAAGAATCGCCTTTAAGTTACAGCAGACGCGCTCAGGAAGCTTCAGGTGGCCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 961 TGGGCAAGTTTTCACACCGGTTGTTGATACCTGGCCAGGAGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1021 AAGCATCTTAGGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGAGACTGCCACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1081 GCCAAACCCGAGCAAGGTCGTCCGCAATCCAGGCTGTGCCGCTTCACTGCGGGAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1141 ACCCTCTGCAAGGGGAGACAGGAGGCCCTTCCAGGCCACAGCCGGTTCCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
Db 1201 ATTTGGTACGTTCCCGCCACGGGCACTCTGTTCTTCCACACCTGTCTGTGGGGCT 1260

QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1321 GGAGGACCTCAGGGTCTCTGAGATTCCTGCAGTGAACCTCCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1381 GACTTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGTTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1441 CCGCCCATATCCAGAGAGACCACTGGCCCTGCTGGAGGCTGAACACCTGCGGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1501 CGGGTGAGATCGGCTTTGCTGTCGACCCAGCACTGATCCGACCAAGGGCCACCTTCAGA 1560
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1561 GCTAGGACCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1621 GAAAGTTCTCAGTTCCTCCAGGCTTC 1647
RESULT 3
AAA15001
ID AAA15001 standard; cDNA; 1993 BP.
XX
AC AAA15001;
XX
DT 21-AUG-2000 (first entry)
XX
DE cDNA encoding a human proliferation and apoptosis related protein.
XX
KW Human; proliferation and apoptosis related protein; PROAP; psoriasis;
KW cell proliferative disorder; immunological disorder; hepatitis;
KW reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma;
KW cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia;
KW asthma; diabetes mellitus; osteoarthritis; endometriosis;
KW uterine fibroid; menstrual cycle; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 42..1691
FT /*tag= a
FT /product= "proliferation and apoptosis related protein"
XX
PN WO200023589-A2.
XX
PD 27-APR-2000.
XX
XX
PF 19-OCT-1999; 99WO-US024511.
XX
XX 20-OCT-1998; 98US-0172216P.
PR 04-FEB-1999; 99US-0118559P.
PR 11-FEB-1999; 99US-0172229P.
PR 22-APR-1999; 99US-0154336P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;
PI Azimzai Y, Baughn MR, Yang J, Shih LL;
XX
XX WPI; 2000-339688/29.
DR P-PSDB; AAY84901.
XX
PT New human proliferation and apoptosis related protein polypeptides used

PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
XX
XX Claim 9; Page 117-118; 128pp; English.
XX
CC The present sequence encodes a human proliferation and apoptosis related
CC protein (PROAP). The polypeptides and polynucleotides can be used for the
CC diagnosis, treatment and prevention of cell proliferative, immunological
CC and reproductive disorders. Disorders associated with decreased
CC expression or activity of include arteriosclerosis, cirrhosis, hepatitis,
CC psoriasis, melanoma, lymphoma and cancers of the breast, brain and
CC prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia,
CC asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine
CC fibroids and disruptions of the menstrual cycle. Antibodies against PROAP
CC can be used in diagnosis of disorders characterized by PROAP e.g. in ELISA
CC to detect and quantify gene expression in biopsied tissues. These
CC techniques can also be used to monitor regulation of PROAP levels during
CC therapeutic intervention
XX
SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.45e-128 Length: 1993
Score: 2978.00 Matches: 548
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.83% Indels: 0
DB: 3 Gaps: 0
US-10-071-838-2 (1-549) x AAA15001 (1-1993)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 42 ATGACCTGCTAGAGGTGCGGGCAGTTGGTGGCACAAGAGGAGAGACATATTATG 101
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 102 AAATACGAAAAGGACACCGAGCTGGCTGCAGAGACAAAGGGCTTAAGCTTTTGA 161
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 162 AGCTACACACACACAGTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTG 221
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluLysSerArgLysSerLysTrpValAsp 80
Db 222 ACTGCGCGGAGCGCAAGCAAAATTCCGGGAGATCAGCCGAAAGACAAAGTGGTGAT 281
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 282 ATGCTGGGAGACTGGAGAAATACAAAAGCAGACAAAGCTCATGATCGAGCTACAG 341
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 342 GAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCGAACAACCTGAGGAATG 401
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 402 AAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAGGGCAAGAGTCTCTGAG 461
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 462 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAAGGAAGCATATTCTTCAGG 521
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 522 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGCGCATATGAGAGTAT 581
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 582 AACCCGAGGTGGCTACTGACGGAACCTGAGCCACATCGCCGCTTCTCTCTAT 641
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

Db 642 CTTCTGAGGAGATGCATTCTGGGCACTGCTGCAGCTGCTGGCCAGTAGAGGCACTCC 701
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGly 240
Db 702 CTGACGGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAGAACAAGACAGAG 761
Qy 241 HisValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 762 CATGTGTAGCCACGTACCAACCCAAAGCAATGGGGCATCAGGACAAAGAAAGATCTATGT 821
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 822 GGGCAGTGTCCCGTTAGGCTGCTCATCCGGAATATGATGACGGGATCTCTCTCGGG 881
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 882 CTCACCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGAAGCAGCGCTGTATGCCGATA 941
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 942 ACAAGATCGCTTTAAGGTTCAAGCAGAAAGCGCTCACGAAGACGTCAGGTGTGGCCCG 1001
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1002 TGGGCACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1061
Qy 341 LysHisLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1062 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1121
Qy 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1122 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTACAGTGGCGGAG 1181
Qy 381 ThrLeuCybLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1182 ACCCTCTGCAAGGGGGAAGGCAAGGCCCTCTCCAGGCCACCAAGCCCGTTCCCGCGGCC 1241
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCybProGlyValAla 420
Db 1242 ATTTGGTACGCTTCCCGCCACGCGCACCTCGTCTTCCACACCTGTGCTGGGGCT 1301
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1302 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCACGCCGCTGTGCTCAG 1361
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1362 GGAGGACCTCAGGGTCTCTGAGATTCTGTGAGTGAATCCCATGCCCGCTCCCAAG 1421
Qy 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1422 GACCTGACGTAAGGGCCCTTGGTTCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1481
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1482 CGTGCCATATCCAGAGAGACAGCTGGCCCCCTGTGGCAGGCTGAACACCCCTGGGAG 1541
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1542 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGCAACCCCTTCAGA 1601
Qy 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCybLeuCybGlyLeuHisLeu 540
Db 1602 GCTAGGAGCAACAGCGGTGTGCTCCACCTCAGGGGCTTGCCTCGGGGCTCCACTTG 1661
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1662 GAAAGTTCTCAGTTCCTCCAGGCTTC 1688

RESULT 4
ADC37384

ID ADC37384 standard; DNA; 1964 BP.
XX AC ADC37384;
XX DT 18-DEC-2003 (first entry)
XX DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 217.
XX XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
XX KW cancer; infectious disease; bone disease; AIDS;
XX KW neurodegenerative disease; ischaemic disorder; Anti-inflammatory;
XX KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
XX KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX OS Homo sapiens.
XX PN WO2003048202-A2.
XX PD 12-JUN-2003.
XX PF 03-DEC-2002; 2002WO-JP012644.
XX PR 03-DEC-2001; 2001JP-00368692.
XX PR 05-DEC-2001; 2001US-0335829P.
XX PR 03-OCT-2002; 2002JP-00291302.
XX PR 04-OCT-2002; 2002US-0415769P.
XX PA (ASAH) ASAH KASEI KK.
XX PI Matsuda A, Muramatsu S;
XX DR MPI; 2003-505282/47.
XX DR P-PSDB; ADC37385.
XX XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX PT useful for treating inflammation, autoimmune diseases, cancers,
XX PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX PT ischemic disorders.
XX PS Claim 4; SEQ ID NO 217; 938bp; English.
XX XX
XX CC The present invention relates to novel proteins and their coding
XX CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
XX CC kappaB). The proteins and their coding sequences are useful for treating
XX CC a disease associated with NF-kappaB activation, such as inflammation,
XX CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
XX CC neurodegenerative diseases, or ischaemic disorders.
XX SQ Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,31e-128 Length: 1964
Score: 2975.00 Matches: 548
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.73% Indels: 0
DB: 10 Gaps: 0

US-10-071-838-2 (1-549) x ADC37384 (1-1964)
Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGACGTGTAGAGGTGCGGCGCAGTTGTGGCACAAAGCAGAGACATCATATG 60
Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AATACGAAAGGGAACAACGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTCA 120
Qy 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 121 AGCTACAAACAACAGTGCATCATTTGGGGAATTGTACATGACAGGAGCTGCTCTTG 180
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

```

Db      181 ACTGCGGGAGGCGGAAGCAAAATTGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
QY      81 MetLeuGlyAspTrpGlyLysSerSerArgLysLeuLeaAspArgAlaTyrLys 100
Db      241 ATGCTGGAGACTGGAGAAATACAAAAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
QY      101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db      301 GGAATGCCCATGACATCCGGGGCCCGATGTGTAGTCTCTGACATTGAGGAAATG 360
QY      121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysGlyLysArgSerSerGlu 140
Db      361 AAGTTGAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
QY      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      421 CACATCCAGCGCATCGACCGGAGCTTAAGCGGACATTAAGGAAGCATATTTCTTCAGG 480
QY      161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db      481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCATCTCTGCGCATATAGAGAGTAT 540
QY      181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db      541 AACC CGAGGTGGGCTACTGCAAGGAGCCTGAGCCATCGCCGCTTGTCTCTCTAT 600
QY      201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db      601 CTTCCTGAGGAGGATGCACTTCTGGCAGCTGTGACAGTCTGCGCCAGTGAAGGCACTCC 660
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db      661 CTGCAAGGATTTCAACGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAGAG 720
QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      721 CATGTGTAGCCACGTCACAACCCAAAGACCATGAGGCAATCAGGACAAAGAAAGATCTATGT 780
QY      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyTyrLeuSerLeuGly 280
Db      781 GGGCAGTGTCCCGCTTGGCTGCTCATCCGGATATGATTGACGGGATCTCTCTGGG 840
QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db      841 CTCACCCCTGGCCTGTGGAGCTGTATCTGTGAAGAGCGCAAGCGCTTGATGCCAATA 900
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db      901 ACAAGATCGCCTTAAGTTCAAGCAAGAGCGCTCAGCAAGAGCGTCCAGGTGTGGCCCG 960
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db      961 TGGGACGTTTTCGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1020
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db      1021 AAGCATCTTAGGGCTCTATGAGAACTAAACAAGAAAGAGGGGAGCTGCAACCCCA 1080
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLys 380
Db      1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1140
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      1141 ACCCTGTGCAAGGGGAGCAGGCAAGGCCCTCCAGGCCCAACAGCGGTTCCCGCGGCC 1200
QY      401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
Db      1201 ATTGTGACGTTCCCGCCACGGGCACTCTGTTCTTCACACCTGTCTGTGGGGCT 1260
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

```

```

Db      1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTAG 1320
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db      1321 GGAGGACTCAGGGTCTCTGGAGATTCTGTGAGTGAAGTCACTGAGTCCCGCTCCCAAG 1380
QY      461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db      1381 GACCTGACGTAGAGGGCCCTGTGTCGCAATTATGATTCAAGACAGAGCTGTGGGTC 1440
QY      481 ArgAlaIleSerGlnGlnLysGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db      1441 CGTGCATATCCAGAGAGACCAAGCTGGCCCTGTGCGAGGCTGAACACCTGCGGAG 1500
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      1501 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCAACCCCTTCAGA 1560
QY      521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLysLeuHisLeu 540
Db      1561 GCTAGGAGCAGACAGCCGTGTCTCTCCACCTCAGGGCTTGCCTGTGGGCTCCACTTG 1620
QY      541 GluSerSerGlnPheProProGlyPhe 549
Db      1621 GAAAGTTCAGTTCCTCCAGGCTTC 1647

RESULT 5
ABX71403
ID   ABX71403 standard; cDNA; 2072 BP.
XX
AC   ABX71403;
XX
DT   14-APR-2003 (first entry)
XX
DE   Human cell cycle-associated cDNA from clone DKFzphes3_35p22.
XX
KW   Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200112659-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-IB001496.
XX
PR   18-AUG-1999; 99US-0149499P.
XX
PR   28-SEP-1999; 99US-0156503P.
XX
PA   (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI   Wiemann B.
XX
DR   WPI; 2001-327840/34.
XX
DR   P-PSDB; ABU53234.
XX
PT   Nucleic acids having the sequences of clones isolated from libraries of
XX
PT   different human tissues, useful in recombinant DNA methodologies.
XX
PS   Claim 1; Page 867; 1095pp; English.
XX
CC   This invention describes novel polynucleotides and polypeptides isolated
CC   from human cDNA libraries which can be used for gene therapy or in
CC   vaccines. The polynucleotides of the invention and antibodies encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate polypeptide expression. The products of the
CC   invention may also be used to identify modulators of expression and
CC   activity and to down regulate expression and activity. The antibodies of
CC   the invention may also be used as diagnostic agents for detecting the
CC   presence of polypeptides in samples. This sequence encodes a polypeptide
CC   described in the disclosure of the invention
XX
SQ   Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	1.24e-127	Length:	2072
Score:	2963.00	Matches:	545
Percent Similarity:	99.64%	Conservative:	2
Best Local Similarity:	99.27%	Mismatches:	2
Query Match:	99.33%	Indels:	0
DB:	5	Gaps:	0

US-10-071-838-2 (1-549) x ABX71403 (1-2072)

QY	1	MecAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
DB	99	ATGACGCTGGTAGAGGTCCGGCGCAGTGTGGGCAAGAAGCGAGAGCATCATATG	158
QY	21	LySerTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	159	AAATAAGAAAGGACACCGAGCTGGGCTGCCAGAGGACAAGGGCCCTAAGCCTTTTCGA	218
QY	41	SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu	60
DB	219	AGCTAACAAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTG	278
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
DB	279	ACTGCGCGGAGCGGAGCAAAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGGTGAT	338
QY	81	MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
DB	339	ATGCTGGAGAGACTGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGCGTACAAG	398
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet	120
DB	399	GGATGCCCCATGAACATCCCGGGCCCGATGTGTCACTCTCTGAACTGAGGAAATG	458
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
DB	459	AACTTGAATAAACCCCGGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCATCTGAG	518
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
DB	519	CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAGCATATATCTTCAGG	578
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr	180
DB	579	GATCGATACGGAACCAAGCAGGGGAATCTCCACATCTCTGCGCATATGAGGAGTAC	638
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr	200
DB	639	AACTCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT	698
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
DB	699	CTTCCTGAGGAGATGATCTGCGCACTGTGCACTGTGCGCAAGTGAAGGCACTCC	758
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
DB	759	CTGCAGGGATTTCACAGCCCAATGCGGGACCGTCCAGGGCTCCAAGACCAACAGAG	818
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
DB	819	CATGTGGTAGCCACGTCACAAACCAAGACCATGCGGCATCAGAGCAAGAAGATCTATGT	878
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
DB	879	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATGACGGGATCTCTCGGG	938
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle	300
DB	939	CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAAAGCGCAACAGCGCTGATCCGATA	998
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320

DB	999	ACAAGAATCGCCTTTAAGGTTCAAGCAGAGCGGCTCAAGAACGTTCCAGGTGTGCCCCG	1058
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
DB	1059	TGGGCAAGTTTGTGCAACCGGTTGTTGATACCTGGGCCAGGATGAGGACATGTGCTC	1118
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
DB	1119	AAGCATCTTAGGGCCCTCTATGAAGAACTAACAGAAAGAGGGGACCTGCCACCCCA	1178
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
DB	1179	GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG	1238
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
DB	1239	ACCTCTGCAAGGGGAGCAGGCGCCCTCCAGGCCACACAGCCCGGTTCCCGGGCCC	1298
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla	420
DB	1299	ATTGTGTCAGCTTCCCGCCACGCGGCACTCGTCTTCCACACCCCTGTGTGGGGCT	1358
QY	421	ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
DB	1359	GTCGGGAAGACACCTACCTGTGGGCACTCAGGCTGTGCCAGCCCGGCTGGCTAG	1418
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
DB	1419	GGAGCACTCAGGGTCTCTGAGATTCTCGCAGTGAACCTCATGCCGCCCTCCCAACG	1478
QY	461	AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
DB	1479	GACTGGAAGTAGAGGGCCCTTGCTCCGCCATTAATGATTACAGACAGAGCTGGGTC	1538
QY	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu	500
DB	1539	CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG	1598
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
DB	1599	CGGGTGAGATCGGCTTGTGGCTGCACCAAGACATGATTCGACAGGCGACCCCTTCAGA	1658
QY	521	AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
DB	1659	GCTAGGGAAGAACAGCAGTGTGCTCCACCTCAGGGCCTTGCTCTGCGGCTCCACTTG	1718
QY	541	GluSerSerGlnPheProProGlyPhe	549
DB	1719	GAAGTTCTCAGTCCCTCCAGGCTTC	1745
RESULT 6			
ABSS4707	ID		
ABSS4707	standard; cDNA; 2146 BP.		
AC	ABSS4707;		
XX	28-NOV-2002 (first entry)		
DT	CDNA encoding human PRC17 protein splice variant 1.		
XX	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;		
KW	breast cancer; cytosstatic; chromosome 17q11-12; gene; ss; splice variant.		
KW	Homo sapiens.		
OS	Location/Qualifiers		
XX	1..1832		
FH	Key		
FT	CDS		
FT	/*tag= a		
FT	/product= "Human PRC17 protein splice variant 1"		
FT	/transl_except= (pos:421..422, aa:Asp)		
FT	/note= "This codon has an apparent 1 nucleotide deletion		
FT	which alters the reading frame"		
XX			

PN WO200262958-A2.
XX 15-AUG-2002.
XX 08-FEB-2002; 2002WO-US003457.
XX 08-FEB-2001; 2001US-0267615P.
XX (TULA-) TULARIK INC.
PI Li J, Powers S, Xiang P, Peng Y;
XX WPI; 2002-706902/76.
DR P-PSDB; ABG70737.
XX
PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
XX or expression relative to normal, for example cancer.
PS Claim 24; Page 63-64; 78pp; English.

CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence encodes the human PRC17 protein
CC splice variant 1 of the invention. The human PRC17 gene is located on
XX chromosome 17q11-12
XX
SQ Sequence 2146 BP; 494 A; 619 C; 616 G; 417 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 4.54e-125 Length: 2146
Percent Similarity: 2907.50 Matches: 547
Best Local Similarity: 89.67% Conservative: 0
Query Match: 97.47% Mismatches: 2
DB: 6 Indels: 62
Gaps: 1

US-10-071-838-2 (1-549) x ABS54707 (1-2146)

QY 1 MetaSpValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGACGTGTAGAGGTGCGGGCAGTTGGTGGCAAGAAGCGAGGACATCATATG 60
QY 21 LysTyrgLulysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 120
QY 41 SerTyraAsnAsnAsnValAspHisLeuGlyIleValHis----- 53
Db 121 AGCTACAACAAACAGTCGATCATTTGGGGATTGTACAGTCTGCCGCTCGGAGTCA 180
QY 53 ----- 53
Db 181 GCCCCACAGGAAGGCCCTTGTCTCCCTTCCCTGTGCTTCTTCTGGGCTGAGCCCTGAG 240
QY 53 ----- 53
Db 241 CTGGAAGGACAGACAGCTCTTCTGTGGGGTCCGCCACCGAGCTGGGGCCGCTCCAG 300
QY 54 ----- 54
Db 301 GCCCCGTGACATTCTCAGCTCTGCTGGTGCCTTACAGTGAACGAGCTGCTCTCT 360
QY 60 LeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79
Db 361 CTGACTGCGGGGAGGCGGAACAATTGCGGGGAGATCAGCCGAAAGAGCAAGTGGGT- 419

QY 80 AspMetLeuGlyAspTrpGluLysTyrsSerSerArgLysLeuIleAspArgAlaTyr 99
Db 420 GATATGCTGGAGACTGGGAGAAATACAAAAGCAGCAAGAAAGCTCATGATCGAGCGTAC 479
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlu 119
Db 480 AAGGGAATGCCCATGAACATCCGGGGCCCGATGTGTAGTCTCTTGAACATTGAGGAA 539
QY 120 MetLysLeuLysAsnProGlyArgTyrgLulMetLysGlyLulysArgSerSer 139
Db 540 ATGAAGTTGAAAACCCTGGAAAGATACCATCATGAAGAGAGAGGCAAGGTCATCT 599
QY 140 GluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159
Db 600 GAGCACATCCAGCGCATCGACCCGGAGCTTAAGCGGACATTAAAGAGCATATATTCTTC 659
QY 160 ArgAspArgTyrgLylThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrgLul 179
Db 660 AGGGATCGATACGGAACCAAGCAGCGGAACTTCCACATCTCTGGCATATGAGGAG 719
QY 180 TyraAsnProGluValGlyTyrgLylArgAspLysSerHisIleAlaAlaLeuPheLeu 199
Db 720 TATTAACCCGGAGTGGCTACTCGACGGAACCTTGAGCCACATCCCGCTTGTCTCTC 779
QY 200 TyrlLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219
Db 780 TATCTTCTGAGAGAGATGATCTTGGGCACTGTGACGTCTGGCCAGTGAAGAGGCAC 839
QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 239
Db 840 TCCCTGACGAGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAG 899
QY 240 GluHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 259
Db 900 GAGCATGTGTAGCCACGTACACCAACCAAGACATGGGGCATCAGACAAGAAAGATCTA 959
QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu 279
Db 960 TGTGGGACAGTTCCTCCGTTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCTC 1019
QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrlLeuValGluGlyGluGlnAlaLeuMetPro 299
Db 1020 GGGCTCACCCCTCGCCTGTGGGACGTGTACTGTGTAGAGCGGAACAGCGCTGATGCCG 1079
QY 300 IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGly 319
Db 1080 ATAAACAAGATCGCCTTTAAGGTTACAGCAAGAGCGCTCACGAAGACCTCAGGTGTGC 1139
QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 339
Db 1140 CCGTGGGACAGTTTGTGCAACCGGTTCTGTGATACCTGGCCAGGATGAGACACTGTG 1199
QY 340 LeuLysHisLeuAlaArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 359
Db 1200 CTCAAGCATCTTAAGGCTCTATGAAGAACTTAACAAAGAAAGAGGAGGACTGCCACCC 1259
QY 360 ProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGly 379
Db 1260 CCAGCCAAACCCAGCAAGGAGTGTGCGGATCCAGGCTGTGCGGCTTCACTGCGGG 1319
QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 399
Db 1320 AAGACCTCTGCAAGGGGAGACAGGAGGCGCTCCAGGCCACAGCCCGGTTCCCGGG 1379
QY 400 ProIleTrpSerAlaSerProProArgAlaProArgSerSerTrpProCysProGlyGly 419
Db 1380 CCCATTTGTCACTTCCCGCCACGGGCACTCTTCTTCCACACCTGTCTGCTGGTGG 1439
QY 420 AlaValArgGluAspThrTyrlProValGlyThrGlnGlyValProSerProAlaLeuAla 439
Db 1440 GCTGTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCT 1499
QY 440 GlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro 459

Db 1500 CAGGAGGACCTCAGGGTTCTTGAGATTCTTCAGTGAACCTCCATCCCGCCCTCCCA 1559
QY 460 ThrAspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 479
Db 1560 ACCGACCTGACGTAGAGGCCCCCTTGTTCCGCATTATGATTTCAGACAGAGCTGCTGG 1619
QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAla 499
Db 1620 GTCCGTGCATATCCCAAGAGACAGCTGCGCCCTGCTGGCAGGCTGAACACCCCTGCG 1679
QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPhe 519
Db 1680 GAGCGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACGAGGCAACCCCTTC 1739
QY 520 ArgAlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHis 539
Db 1740 AGAGCTAGGAGCGAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCAC 1799
QY 540 LeuGluSerSerGlnPheProGlyPhe 549
Db 1800 TTGGAAAGTTCTCAGTTCCTCCAGGCTTC 1829

RESULT 7

ADC37380 ID ADC37380 standard; DNA; 2304 BP.

AC ADC37380;

DT 18-DEC-2003 (first entry)

XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neutropoietic; Nootropic; Cardiant; Gene therapy; human; gene; ds.

OS Homo sapiens.

PN WO2003048202-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

XX (ASAH) ASAH KASEI KK.

XX Matsuda A, Muramatsu S;

DR WPI; 2003-505282/47.

DR P-PSDB; ADC37381.

XX New purified protein that activates nuclear factor kappa B (NF-kappaB),

PT useful for treating inflammation, autoimmune diseases, cancers,

PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

PT ischaemic disorders.

XX Claim 4; SEQ ID NO 213; 938bp; English.

XX The present invention relates to novel proteins and their coding

CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-

CC kappaB). The proteins and their coding sequences are useful for treating

CC a disease associated with NF-kappaB activation, such as inflammation,

CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,

CC neurodegenerative diseases, or ischaemic disorders.

SQ Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.4e-125 Length: 2304
Score: 2906.50 Matches: 540
Percent Similarity: 98.36% Conservative: 0
Best Local Similarity: 98.36% Mismatches: 2
Query Match: 97.44% Indels: 8
DB: 10 Gaps: 1

US-10-071-838-2 (1-549) x ADC37380 (1-2304)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 362 ATGACGTGTAGAGGTCCGGGCGAGTTGTGGCCACAAGACCGAGAGACATCATTTATG 421
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 422 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 481
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 482 AGCTACACACAAACGTCGATCATTTGGGATGTATCATGAGACGAGCTGCCTCTCTG 541
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 542 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAGACCAAGTGGTGAT 601
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 602 ATGCTGGAGACTGGGAGAAATACAAAGCAGACAAAGCTCATGATCGAGCGTACAG 661
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 662 GGAATGCCCATGAAATCCGGGGCCGATGTGTAGTCTCTCTGAACTGAGGAAATG 721
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 722 AAGTTGAAAAACCCCGAAGATACCAATCATGAAAGAGAGAGGCAAGGTCTATCTGAG 781
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 782 CACATCCAGCGCATCGACCGGAGCTAGCGGAGCATTAAGAACATATATCTTCAG 841
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 842 GATCGATACGGAACCAAGCAGCGGGAATCTCCACATCTCTCTGCAATATGAGAGATAT 901
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 902 AACCCGAGGTGGCTACTGACAGGACTGAGCCACATCGCCGCTGTCTCTCTAT 961
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 962 CTTCCTGAGAGAGATGATTTCTGGGCACTGTGACGTGTGGCAGTAGAGGCACTCC 1021
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1022 CTGCAGGATTTCAAGCCCAATGCGGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 1081
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 1082 CATGTGTAGCCACGTACAAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATGT 1141
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1142 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTGGG 1201
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 1202 CTCACCTGCGCTGTGGAGCTGTATCTGTAGAAAGCGAAGCGGTGATGCCGATA 1261
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

DB 1262 ACAGAAATCGCCTTTAAAGGTTGAGCAGAGCGCCTCAGCAAGACGTCCAGGTGTGCCCG 1321
QY 321 TRPALAArgPheCySAAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1322 TGGGCACTTTTTCGAAACCGGTTCTGTATACCTGGGCCAGGGATGAGGACACTGTGCTC 1381
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1382 AAGCATCTTAGGGCCTCTATGAGAAACTAACAAGAAAGCAGGGGGAAGCTGCCACCCCA 1441
QY 361 AlalysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1442
DB 1442 GCCAAACCCGAGCAGGAGGTCGTCCGATCCAGGCTGTGCCGCTTCACTGGCGGGAAG 1501
QY 381 ThrLeuCyLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1502 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCCAGGC-----CCA 1540
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerSerThrProCybProGlyGlyAla 420
DB 1541 TT-TGGTCAGCTTCCCGCCAGCAGGCACTGTTCTTCACACCCCTGCTGTGGGCT 1599
QY 421 ValArgGluAspThrThrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1600 GTCCGGGAAGACACTACCTGTGGCACTCAGGGTGTGCCAGCCGCCCTGCTCAG 1659
QY 441 GlyGlyProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1660
DB 1660 GGAGGACCTCAGGGCTTCCGTGAGATTCTGTGAGAACTCCATGCCCCCGCTCCCAACG 1719
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1720 GACCTGGACGTAGAGGGCCCTTGCTTCCGCCATTATGATTTCAGACAGAGCTGCTGGGTC 1779
QY 481 ArgAlaIleSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1780
DB 1780 CGTCCATATCCACAGAGGAGCAGCTGGGCCCTGTGTCAGGGCTGAACACCTGCGGAG 1839
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1840 CGGGTGAGATCGGCTTTCGTCGACCCAGCACTGATTCCGACAGGAGCCCTTCAGA 1899
QY 521 AlaArgAspGlnGlnProCybAlaProThrSerGlyProCysLeuCybGlyLeuHisLeu 540
DB 1900 GCTAGGAGCAGACGCGGTGTCTCCCACTCAGGGCCTTGCTGCGGCTCCACTTG 1959
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1960 GAAAGTTCAGTTCCTCCAGGCTTC 1986

RESULT 8
ID ADC37378 standard, DNA; 2304 BP.
XX
AC ADC37378;
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matouda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PsDB; ADC37379.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 211; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
SQ Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.26e-124 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 3
Query Match: 97.17% Indels: 8
DB: 10 Gaps: 1

US-10-071-838-2 (1-549) x ADC37378 (1-2304)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 362 ATGACGTGTGTAGAGTCCGCGGCACTTGTGGCACAAGAGCGAGACATCATATG 421
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 422 AATACGAAGAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCCTAAGCCTTTTCCA 481
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 482 AGCTACAACAACAACGTGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 541
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 542 ACTGCGCGGAGAGCGAAGCAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT 601
QY 81 MetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 602 ATGCTGGAGACTGGGAGAATAATCAAAAAGCAGCAAAAGCTCATAGATCGAGCTACAG 661
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 662 GGAATGCCCATGAAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGGAATG 721
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140
DB 722 AAGTTGAATAAACCCTCGAAGATACAGATCATGAAGAGAGGCAAGAGTCTGAG 781
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 782 CACATCCAGCGCATGACCGGGAGCGGACGTAAGCGGAGACATTAAAGAAACATATCTCAGG 841
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180

Db 842 GATCGATACGGAAACCAAGCGGGAACTACTCCATCTCTGGCATATGAGAGTAT 901
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuTyr 200
Db 902 AACCCGAGGTGGGCTTACTGACAGGACCTGAGCCATCGCCCTTGTTCTCTCTAT 961
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 962 CTTCCTGAGAGATGATCATTTCTGGCAGTGTGACCTGTGCTGACAGAGGCACTCC 1021
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1022 CTGCAGGATTTACAGACCCCAATGCGGGACCGTCCAGGGCTCCAGAACCAACAGAG 1081
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 1082 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAGAAAGATCTATGT 1141
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1142 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTCGGG 1201
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 1202 CTCACCCCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 1261
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1262 ACAAGAATCGCCTTTAAGTTACAGCAGAAAGCGCCTCACGAAGACGTCAGGTGTGCCCG 1321
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1322 TGGGCACGTTTGTGCAACGGTTCGTTGATACCTGGGCCAAGGATGAGACACTGTGCTC 1381
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1382 AAGCATCTTAGGGCCCTTATGAAGAACTAACAGAAAGCAGGGGAGCTGCAACCCCA 1441
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1442 GCCAAACCCGAGCAAGGCTCGTCGCATCCAGGCTGTGCGGCTTCACTGCGGGAG 1501
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1502 ACCCTCTGCAAGGGGACAGGCAAGCCCTCCAGGC-----CCA 1540
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1541 TT-TGGTACGCTTCCCCCAAGGGCAGGCACTGCTTCCACACCCCTGCTGCTGCTGCT 1599
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1600 GTCCGGGAAGACACTTACCTGTGGCAGTCCAGGGTGTGCCAGGCCGCCCTGCTCAG 1659
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1660 GGAGGACCTCAGGGGTTCTCTGAGATTCTCTGACGTGAACCTCCATGCCCTCCCAACG 1719
Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1720 GACCTGACGTAGAGGGCCCTTGGTTCGCCCATATGATTTACAGACAGAGCTGCTGGGTC 1779
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1780 CGTGCCATATCCAGAGGAGACAGCTGAGCCCTGCTGGCAGAGCTGAACACCTGCGAG 1839
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1840 CCGGTGAGATCGGCTTTCGCTGACCCAGACACTGATTCGACCAAGGCAAGCCCTTCA 1899
Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540

Db 1900 GCTAGGACGAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTTGGCGGCTCCACTTG 1959
Qy 541 GluSerSerGlnPheProGlyPhe 549
Db 1960 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1986
RESULT 9
ABZ11810
ID ABZ11810 standard; cDNA; 2084 BP.
XX
AC ABZ11810;
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 692.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytoskeletal; immunomodulator; nootropic; neuroprotective;
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR P-PSDB; ABP69593.
XX
XX
PT New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative,
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 692; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
nucleotide sequence selected from any of 948 sequences (ABZ1119-
ABZ12066) or their mature protein coding portion, active domain coding
protein or complementary sequences. The polynucleotides are useful for
identifying expressed genes or for physical mapping of human genome. The
encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
markers, as a food supplement, for generating antibodies, in medical
imaging, screening and diagnostic assays and for treating cell-
proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
platelet or coagulation disorders, wound, burns, incision, ulcers, liver
or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
arthritis, etc. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;
Alignment Scores: 1.94e-122 Length: 2084
Pred. No.: 2850.00 Matches: 528
Score:

Percent Similarity: 96.36%
Best Local Similarity: 96.17%
Query Match: 95.54%
DB: 6
Conservative: 1
Matches: 2
Indels: 18
Gaps: 1

US-10-071-838-2 (1-549) x ABZ11810 (1-2084)

QY 1. MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 163 ATGACGTGTAGAGGTGCGGGCAGTTGGTGGCACAAAGAGCGAGACATCATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AAATACGAAAGGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCGCTTAACCTTTTCA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 283 AGCTACAAACAACAAGTCGATCATTTGGGATTTGATGATGAGACGAGCTGCCCTCTTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 343 ACTGCGGGGAGCGGAAGCAAATTCGGCGGAGATCAGCCGAAAGCAAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 403 ATGCTGGGAGACTGGGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 463 GGAATGCCCATGAACATCCGGGGCCGATGTGTAGTCTCTCTGAACACTGAGGAAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 529 AACATCCAGCGCATCGACCGGAGCTTAAGCGGCACATTAAGGAAGCATATATTCTTCAG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 589 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGCGCATATGAGAGATAT 648
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
DB 649 AACCCGAGGTGGGCTACTGCGAGGACTGAGCCACATCGCCGCTTGTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 709 CTTCTGAGAGAGATGATTCATTCGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 769 CTGCAAGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGACCAACAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 829 CATGTGTAGCCACGTCACAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 889 GGGCAGTGTCCCGTTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCCGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
DB 949 CTCACCCCTGCGCTGTGGGAGCTGTATCTGTAGAAAGCGCAACAGGCGTGTATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 1009 ACAAGAATCGCTTTAAGTTCAAGCAGAAGCGCTCAGCAAGACGTCCAGGTGTGCGCCG 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340

DB 1069 TGGGCACGTTTTTGCACACCGGTTCCGTTGATACCTGGGCCAGGATGAGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1129 AAGCATCTTAGGGCCTCTATGAAGAATACTAACAAGAAAGAGGGGAGCTTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1189 GCCAAACCCGAGCAAGGTCGTGCGCATCAGGCCCTGTGCGGCTTCACTGCGGGAAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1249 ACCCTCTGCAAGGGGACAGCAGGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1309 ATTTGGTCAGCTTCCCGCCACAGGGACCTCGTCTTCCACACACCTGTCTGTGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1369 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCCAAGCCCGGCTGGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1429 GAGGACCTCAGGGTTCCTGAGATTCTGCAGTGGAATCTCATGCCCGCTCCCAAGC 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1489 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGGT 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1549 CGTGCCATATCCAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAAACACCTGCGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1609 CGGGTGAGATCGGCTTTCGTGCAACCAAGCACTGATTCGACCAAGGGCACTGAGAG 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1669 GCTAGGAGCAGAACAGCGGTGTGCTCCCACTCAGGGCCTTGCTGTGCGGCTCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1729 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755

RESULT 10
ADM44328
ID ADM44328 standard; cDNA, 2084 BP.
XX
AC ADM44328;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #692.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.
PR 20-AUG-2002; 2002US-00225251.
XX
PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.

PA (DRMA/) DRMANAC R T.
XX XX
PI Tang YT, Xue A, Drmanac RT;
XX
DR MPI; 2004-238579/22.

XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.

XX Disclosure; SEQ ID NO 692; 51pp; English.

XX The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.

XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.94e-122 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.36% Conservative: 1
Best Local Similarity: 96.17% Mismatches: 2
Query Match: 95.54% Indels: 18
DB: 12 Gaps: 1

US-10-071-838-2 (1-549) x ADM44328 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGCTGTAGAGGTCCGCGGCAGTGTGTGGCAACAAGAGAGAGATCATATTATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCCTTAAGCTTTTGA 282
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACAAACAACACGTCATCTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGCGGAGCAAAATTCGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGACGAAAGCTCATAGATGACGCTACAAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGAGCGTAAGCGGAGACATTAAAGAACATATATTCTTCAGG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGGAACCAAGACCGGAACTACTCCATCTCTGCGCATATGAGAGATAT 648
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGCTACTGACGAGGACCTGAGCCACATCGCGCTGTCTCTCTCTAT 708

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCCTGAGGAGATGATCTTGGGCACTGGTGCACTGCTGCGCCAGAGAGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGACGGATTTCACAGCCCAAAATGGCGGACCGTCCAGGGGCTCCAGAACCAACAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTCCCGGTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 949 CTCACCTGCGCCTGTGGAGCGTGTATCTGTAGAAAGCGAACAGCGCTTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCCTTTAAGTTTACAGCAAGAGCCCTCACGAAGACCTTCAGGTGTGCCCG 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGCACGTTTGTGCAACCGTTCTGTGATACCTGGCCAGGAGTGAAGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGCGCTTATGAAGAACTAACAAAGACAGGGGACCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1189 GCCAAACCCGAGCAAGGCTGTCGGCATCCAGGCTGTGCGGCTTCACTGCGGGAAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGCAGGACAGGCCCTCTCAAGGCCACCAAGCCCGGCTTCCGCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
Db 1309 ATTGTGACGTTCCCGCCACCGGACCTGCTTCTTCCACACACCTGTCTGTGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACACCTACCTGTGGCACTCAGGGTGTGCCAGCCCGCCTGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GAGAGACCTCAGGGTCTCTGAGATTCTGTGAGTGAAGTCCATGCCCGCTCCCAACG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGACGTAAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 1549 CGTGCCATATCCAGAGAGACAGCTGGCCCTGTGTCAGAGCTGAACACCTGCGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCAACCCCTTCAGA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1669 GCTAGGAGCAAGACCGGTGTGTCTCCACTCAGGGCCTTGCTGCGGCTCCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1729 GAAAGTTCTCAGTTCCTCCAGGCTTC 1755

QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1949 GTCCGGAGACACACTCACTCTGTGGCACATCGAGGTGTGCCACAGCCCGCCTGTGCTAG 2008
QY 441 GLYGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 2009 GGAGGACCTCAGGGTTCCTGGAGATTCTCGAGTGAAGTCCATGCCCCGCTCCCAAG 2068
QY 461 AspleuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 2069 GACCTGGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
DB 2129 CGTGCCATATCCAGAGGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGGCGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 2189 CCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACACAGGGCACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 2249 GCTAGGAGCAGACAGCGGTGTGTCCACCTCAGGGCCTTGCTGTGGGGCCTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 2309 GAAAGTCTCAGTTCCTCCAGGCTTC 2335

RESULT 12

AB211809
ID AB211809 standard; cDNA; 2072 BP.

XX AB211809;

XX 20-JAN-2003 (first entry)

DE Human polynucleotide SEQ ID NO 691.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI: 2002-759812/82.

DR P-P8DB; ABP69592.

XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.

XX Claim 1; SEQ ID NO 691; 1012pp + Sequence listing; English.

CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB211119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP6949) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,19e-121 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x AB211809 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 163 ATGAGCGTGGTAGAGGTTCGCGGCAGTTGTGGGCACAAAGAGCGAGACATCATTTATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AAATACGAAGAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 282
QY 41 SerTyrAsnAsnAsnValAspHisIleLeuGlyValHisGluThrGluLeuProProLeu 60
DB 283 AGCTACAAACAACAGCTGATCATTTGGGATTTGATGAGACGAGCTGCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 343 ACTGCGCGGAGGGAAGCAAAATTCGGCGGAGATCAGCCGAAGAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 403 ATGCTGGAGACTGGAGAAATACAAAGACAGAAAGCTCATAGATCGAGCTACAAAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 463 GGAATGCCCATGAAATATCCGGGGCCGATGTGTCACTCCTCGAACAACACTGAGAAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 523 AAGTTGAAAAACCCCGGAAGATACCATCATGAAGAGAGGCAAGATCATCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 583 CACATCCAGCGCATTCGACCGGACGTAAAGCGGACATTAAAGACATATATTTCTTCAG 642
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 643 GATCGATACGGAACCAAGCAGCGGAATCACTCCACATCTCTGACATATGAGAGATAT 702
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
DB 703 AACCCGAGGTGGCTACTGACGAGACTGAGCCACATCGCGCCTTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 763 CTTCCTGAGAGATGATCTTCTGGCACTGTGACGCTGCTGGCCAGTGAAGAGCACTCC 822

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 823 CTGCAGGAGATTTCACAGCCCAATGGCGGAGACCGTCCAGGGGGCTCCAGAACCAACAGAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuGly 260
DB 883 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCAATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrIleuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 937 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAGCGCAAGCGGCTTGATGCCGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 997 ACAGAATCGCCTTTAAGGTTACAGACAGCGCTCACGAAGACGTCACAGGTGTGGCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1057 TGGGCACGTTTTCGAAACGGTTCGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1117 AAGCATCTTAGGGCTCTATGAAAGAACTAACAAGAAAGCAGGGGACCTGCAACCCCA 1176
QY 361 AlaLysProGlnGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1177 GCCAACCAGCAAGGAGGTGTCGACATCCAGGCGCTGTGCCGCTTCACTGTGGGGAAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1237 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCCAGGCCACACGCCGCTTCCCGGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1297 ATTGTCAGCTTCCCGCCACGCGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1357 GTCCGGGAGACACTAACCCTGTGGGCACTCAGGGGTGCCCCAGCCCGGCTGTGGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1417 GGAGGACCTCAGGCTTCTGAGATTCCTGAGTGAATCCATGCCCCCTCCCAACG 1476
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1477 GACCTGACGTAGAGGGCCCTTGGTTCGCCCATATGATTTCAGACAGAGCTGTGGGTC 1536
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1537 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGCTGCAGGCTGAACACCTTGGGAG 1596
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1597 CGGGTGAATCGGCTTTCGCTGACCCAGCACTGATTCGACCAAGGCAACCCCTTCAAGA 1656
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
DB 1657 GCTAGGAGCAGACAGCGGTGTCTCCCACTCAGGGGCTTGCCTCTCGGCTTCACTTG 1716
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1717 GAAAGTTCTCAGTTCCCTCAGGCTTC 1743
RESULT 13
ADM44327
ID ADM44327 standard; cDNA; 2072 BP.
XX
AC ADM44327;

XX 03-JUN-2004 (first entry)
DT
XX
DE Novel human arginine-rich protein cDNA #691.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
XX genetic disorder.
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.
PR 20-AUG-2002; 2002US-00225251.
XX
PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
PI Tang YT, Xue A, Drmanac RT;
XX
DR WPI; 2004-238579/22.
XX
PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
PS Disclosure; SEQ ID NO 691; 51bp; English.
XX
SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2,19e-121 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
DB: 12 Gaps: 1
US-10-071-838-2 (1-549) x ADM44327 (1-2072)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 163 ATGACGTGTAGAGGTGCGGGCAGATTGTTGGGCACAAAGCGAGAGCATCATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AATATACGAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGGCTTAAGCCTTTTGA 282
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGlnLeuProProLeu 60
DB 283 AGCTACAAACAACAGTCGATCATTTGGGATTTGATCATGAGCGAGCTGCTCTTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 343 ACTGCGCGGAGCGGAAGCAATTCCGCCGGAGATCAGCCGAAAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100

Db 403 ATGCTGGAGACTGGAGAAATACAAAAGCAGAAAGCTCATAGATCGAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGGTCAGTCTCTGTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrglnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTGAAAAACCCCGAAGATACACAGATCATGAAGAGAGGGCAGAGGTCATCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 583 CACATCCAGCGCATCGACCGGAGCGTAAGCGGGACATTAAAGAGCATATATTCTTCAG 642
QY 161 AspArgTyrglyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrgluGlyTy 180
Db 643 GATCGATACGGAAACCAAGCAGCGGAACTACTCCACATCTCTGCGCATATGAGAGTAT 702
QY 181 AsnProGluValGlyTyrglyCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTy 200
Db 703 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 220
Db 763 CTTCCTGAGGAGGATGATCTTCTGGCACGTGGTGCAGCTGCTGGCAGTAGAGGCACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 823 CTGCAGGATTTCAAGCCCAATATGGGGACCGTCCAGGGGCTCCAAAGCAACAGAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 883 CATGTGTAAGCCAGCTACAACCCAGACCATGGGGCATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrlleuValGluGlyGlnAlaLeuMetProIle 300
Db 937 CTCACCTGCGCTGTGGAGCGTGTATCTGTGAAGGCGAAGCGGTGATGCGGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGAATCGCCTTAAGTTACAGACAGAGCGCTCACGAAGACGTCCAGGTGTGCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGATGAGGACACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1117 AAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAAGCAGGGGACCTGCAACCCCA 1176
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1177 GCCAAACCCGAGCAAGGGTGTGGGCATCCAGGCTGTGCCGCTTCACTGTGGCGGAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTCTGCAAGGGGACAGGCAAGCGCCCTCCAGGCCACCAAGCCCGGTCCCGCGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTTGTAGCTTCCCCCGCACCGGCACTGTTCTTCACACACCTGTCTGTGGGGCT 1356
QY 421 ValArgGluAspThrTyrrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGGAAGACACTACCTGTGGGCACTCAAGGCTGTGCCCAAGCCCGGCTGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460

Db 1417 GGAGGACCTCAGGGTTCCTGGAGATTCTGCACTGGAAGCTCCATGCCCGCTCCCAAG 1476
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrrAspPheArgGlnSerCysTrpVal 480
Db 1477 GACCTGACGTAGAGGGCCCTTGGTTCCGCATTATGATTTCAAGACAGAGCTGTGGTC 1536
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1537 CGTGCCATATCCAGAGAGACCAAGTGGCCCCCTGTGGCAGGCTGAACACCCCTGGAG 1596
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1597 CCGGTGAGATCGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGGCACCCTTCAGA 1656
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1657 GCTAGGACGAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTGCGGCTCCTG 1716
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1717 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1743
RESULT 14
ADC37386
ID ADC37386 standard; DNA; 2647 BP.
XX
AC ADC37386;
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 219.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI, 2003-505282/47.
DR P-PSDB; ADC37387.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 219; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
SQ Sequence 2647 BP; 609 A; 765 C; 750 G; 523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.76e-121
Score: 2827.00
Percent Similarity: 95.63%
Best Local Similarity: 95.63%
Query Match: 94.77%
DB: 10
Length: 2647
Matches: 525
Conservative: 0
Mismatch: 2
Indels: 22
Gaps: 1

US-10-071-838-2 (1-549) x ADC37386 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 755 ATGACGTGTAGAGGTCCGGCGAGTTGGTGGGCAAGAAGCGAGAGGACATCATATG 814
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 815 AAATACGAAAGGAGACACCGAGCTGGGCTGCCAGAGGACAGGGGCTTAAGCTTTTCGA 874
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlyThrGluLeuProPheLeu 60
DB 875 AGCTACACACACACAGCTCGATCATTTGGGGAATTGATGATGACGAGCTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 935 ACTGCGCGGAGCGGACCAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT 994
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 995 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 1055 GGAATGCCCATGAAATCCGGGCGGATGTGTCTCTGCACTGCAACACTGAGAAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysArgSerSerGlu 140
DB 1115 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGCAAGAGGTCTATGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 1175 CACATCCAGCGCATCGACCGGAGCTAGCGGACATTAAGAAAGCATATATTCTTCAG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGCAATGAGAGTAT 1294
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
DB 1295 AACCCGAGGTGGCTACTGACGGAGCTGAGCCACATCGCCCTTGTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 1355 CTTCTGAGAGGATGATCATCTGGGCACTGGTGCAGCTGCTGCCCAAGTGAAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 1415 CTGACGGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGACCAAGAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 1475 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCACTAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 1517 -----ATCTCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 1529 CTCACCTGCGCTGTGGAGCTGTATCTGTAGAAGCGCAAGCGCTGATGCCGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

DB 1589 ACAAGATCGCCTTTAAAGTTTCAGCAGAAAGCGCTTCACGAGACCTTCAGGTGTGGCCG 1648
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1649 TGGGACAGTTTTCACACCGGTTGTTGATATCTGGGCCAGGATGAGGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1709 AAGCATCTTAGGGCTCTATGAAAGAACTACAAAGAAAGAGGGGGAAGCTGCAACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1769 GCCAAACCGAGCAAGGTCGTCCGATTCAGGCTGTCCGGCTTCACTGGCGGGAAG 1828
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1829 ACCCTCTGCAAGGGGAGACAGGAGGCCCTTCAGGCCCAACAGCCGGTTCCCGCGGCC 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
DB 1889 ATTGTGACGTTCCCGCCAGCGGCACCTCTCTTCCACACCTGTCTGTGGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1949 GTCCGGGAAGACACCTACCTCGTGGGACATCAGGGTGTGCCAGCCCGGCTGCTCAG 2008
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 2009 GAGGACCTCAGGTTCTCGAGATTCCTGCACTGCACTCCAGCCCGCTCCCAAG 2068
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 2069 GACTCGACGTAGAGGCGCTTGTGTTCCGCATTAATTTCAGACAGAGCTGTGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 2129 CGTGCATATCCAGAGAGACCAAGCTGGCCCTGCTGGCAGGCTGAACACCTCGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 2189 CGGGTAGATCGCTTTCCTGACCCAGCACTGATTCGACCAAGGCAAGCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 2249 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCTTGCTGCGGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335
RESULT 15
ADM01991
ID ADM01991 standard; cDNA; 2647 BP.
XX
AC ADM01991;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:676.
XX
KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
OS
PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ieogat T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX WPI, 2003-723558/69.
DR P-PSDB; ADM04434.

XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 676; 305bp; English.

CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM0316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.

XX Sequence 2647 BP; 609 A; 765 C; 750 G; 523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.76e-121	Length:	2647
Score:	2827.00	Matches:	525
Percent Similarity:	95.63%	Conservative:	0
Best Local Similarity:	95.63%	Mismatches:	2
Query Match:	94.77%	Indels:	22
DB:	11	Gaps:	1

US-10-071-838-2 (1-549) x ADM01991 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 755 ATGACGTGTGTAGAGSTCCGGGCACTTGTTGGGCACAAAGACCGAGACATCATTTATG 814
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 815 AAATAGCAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 874
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 875 AGCTACAACAACAACGTGATCATTTGGGGATGTGTACATGAGACGAGCTGCCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 935 ACTGCGCGGAGCGGAGCAACAATTCCGCGGAGATCAGCCGAAGAACAAGTGCGTGAT 994
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 995 ATGCTGGAGACTCGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGAGCGTACAAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 1055 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACACTGAGGAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 1115 AAGTTGAAAAACCCCGAAGATACACAGATCATGAAGAGAGGCAAGGTCTATCGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 1175 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAGCATATATTCTTCAGG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 1235 GATCGATACGGAACCAAGCAGCGGGAATCTCCACATCTCTGCGCATATGAGGAGTAT 1294

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
DB 1295 AACCCGAGGTGGGTACTGACAGGAGACTGAGCCACATCGCCGCTTGTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 1355 CTTCCTGAGGAGATGATCTTCTGGGCACTGTGTCAGCTGTGGCCAGTGAGAGGACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGluAspGlnGln 240
DB 1415 CTGCAGGATTTCAACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 260
DB 1475 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCATCAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 1517 -----ATCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 1529 CTCACCTGCGCTGTGGACGTGTCTGTGTAGAAAGCGAACAGCGTTGATGCCGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 1589 ACAAGAATCGCCTTAAAGTTACAGCAGAACGCCCTCACGAAGCGTCCAGTGTGGCCG 1648
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1649 TGGGCACGTTTGTCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1709 AAGCATCTTAGGCCCTTATAGAAACTTAACAAGAAAGCAGGGGGAACCTGCACACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1769 GCCAAACCCGAGCAAGGTCGTGGCATTCAGGCCCTGTGCGGCTTCACGTGCGGGAAG 1828
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1829 ACCCTGTCAAGGGGAGCAGCAGGCCCCCTCCAGGCCACCAAGCCGGTTCGCCGCC 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1889 ATTTGGTACGTTCCCGCCAGCGGCACTGTTCTTCCACACCTGCTGCTGGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1949 GTCCGGGAAGACACTACCTGTGTGGCACTCAAGGTGTGCCAGCCCGGCTGCTCAG 2008
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 2009 GGAGGACCTCAGGGTCTCTGAGATTCTGAGATTCAGTGAACCTCATGCCCGCTCCCAAG 2068
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 2069 GACCTGACGTAGAGGGCCCTTGTGTCGCCATTATGATTTCAGACAGAGCTGCTGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 2129 CGTGCCATATCCAGAGAGACCAAGCTGCCCCCTGTGCGAGGCTGAACACCTGCGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 2189 CGGGTAGATCGGCTTTCGTGACCAACCACTGATTCCGACCAAGGCAACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 2249 GCTAGGAGCAACAGCCGTGTGTCCCACTCAGGGCCCTTGTGCTGCGGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549

Mon Feb 7 06:58:10 2005

us-10-071-838-2.rng

Page 20

Db 2309 GAAAGTTCTCAGTTCCCTCCAGGCTTC 2335

Search completed: February 4, 2005, 08:20:27
Job time : 819 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 10:14:14 ; Search time 853 Seconds
(without alignments)
3698.116 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983
Sequence: 1 MDVVEVAGSWAQRERDII.....TSGPCLGLHLESSQFPFGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4300275 segs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10071838/runat_03022005_071221_304/app_query.fasta_1.711
-DB=Published Applications NA -QFMT=faaBap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10071838@CGN_1_1_627@runat_03022005_071221_304
-NCPUS=6 -ICPU=3 -NO MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	14 US-10-071-838-1	Sequence 1, Appli
2	2978	99.8	1993	18 US-10-839-882-31	Sequence 31, Appli
3	2907.5	97.5	2146	14 US-10-071-838-3	Sequence 3, Appli
4	2850	95.5	2084	16 US-10-302-172-692	Sequence 692, App
5	2827	94.8	2072	16 US-10-302-172-691	Sequence 691, App
6	2827	94.8	2647	16 US-10-108-260A-676	Sequence 676, App
7	2784	93.3	1752	15 US-10-094-466-61	Sequence 61, Appli
8	2773	93.0	1862	14 US-10-071-838-5	Sequence 5, Appli
9	2191.5	73.5	7878	9 US-09-962-436-562	Sequence 562, App
10	2133.5	71.5	8201	17 US-10-283-975A-272	Sequence 272, App
11	2133.5	71.5	8408	13 US-10-098-841-51	Sequence 51, Appli
12	2133.5	71.5	8420	13 US-10-098-841-50	Sequence 50, Appli
13	1507.5	50.5	9805	9 US-09-764-869-1824	Sequence 1824, Ap
14	1507.5	50.5	9805	10 US-09-764-891-7857	Sequence 7857, Ap
15	1507.5	50.5	9805	14 US-10-091-504-1824	Sequence 1824, Ap
16	1507.5	50.5	9805	16 US-10-227-577-1824	Sequence 1824, Ap
17	1500.5	50.3	9404	9 US-09-764-869-1826	Sequence 1826, Ap
18	1500.5	50.3	9404	10 US-09-764-891-7859	Sequence 7859, Ap
19	1500.5	50.3	9404	14 US-10-091-504-1826	Sequence 1826, Ap
20	1500.5	50.3	9404	16 US-10-227-577-1826	Sequence 1826, Ap
21	1469	49.2	9804	9 US-09-764-869-1829	Sequence 1829, Ap
22	1469	49.2	9804	10 US-09-764-891-7861	Sequence 7861, Ap
23	1469	49.2	9804	14 US-10-091-504-1829	Sequence 1829, Ap
24	1469	49.2	9804	16 US-10-227-577-1829	Sequence 1829, Ap
25	1465	49.1	9796	9 US-09-764-869-1825	Sequence 1825, Ap
26	1465	49.1	9796	10 US-09-764-891-7858	Sequence 7858, Ap
27	1465	49.1	9796	14 US-10-091-504-1825	Sequence 1825, Ap
28	1465	49.1	9796	16 US-10-227-577-1825	Sequence 1825, Ap
29	1428.5	47.9	6696	9 US-09-764-869-1821	Sequence 1821, Ap
30	1428.5	47.9	6696	10 US-09-764-891-7856	Sequence 7856, Ap
31	1428.5	47.9	6696	14 US-10-091-504-1821	Sequence 1821, Ap
32	1428.5	47.9	6696	16 US-10-227-577-1821	Sequence 1821, Ap
33	1355.5	45.4	9831	9 US-09-764-869-1827	Sequence 1827, Ap
34	1355.5	45.4	9831	14 US-10-091-504-1827	Sequence 1827, Ap
35	1355.5	45.4	9831	16 US-10-227-577-1827	Sequence 1827, Ap
36	1355	45.4	3979	9 US-09-764-869-1828	Sequence 1828, Ap
37	1355	45.4	3979	14 US-10-091-504-1828	Sequence 1828, Ap
38	1355	45.4	3979	16 US-10-227-577-1828	Sequence 1828, Ap
39	1053	35.3	886	9 US-09-764-869-1831	Sequence 1831, Ap
40	1053	35.3	886	9 US-09-764-869-1832	Sequence 1832, Ap
41	1053	35.3	886	14 US-10-091-504-1831	Sequence 1831, Ap
42	1053	35.3	886	14 US-10-091-504-1832	Sequence 1832, Ap
43	1053	35.3	886	16 US-10-227-577-1831	Sequence 1831, Ap
44	1053	35.3	886	16 US-10-227-577-1832	Sequence 1832, Ap
45	818	27.4	470	9 US-09-815-343-501	Sequence 501, App

ALIGNMENTS

US-10-071-838-1
; Sequence 1, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: DNA

TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
FILE REFERENCE: PF-0619 PCT
CURRENT APPLICATION NUMBER: US/10/839,882
CURRENT FILING DATE: 2004-05-05
PRIOR APPLICATION NUMBER: US/09/807,452
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/175,737; unassigned; 60/118,559; 09/249,740; unassigned;
60/154,336
PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
1999-04-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PIRL Program
SEQ ID NO: 61
LENGTH: 1993
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 4325626CB1
US-10-839-882-31

54
231
10262

Alignment Scores:
Pred. No.: 3,19e-265 Length: 1993
Score: 2978.00 Matches: 548
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.83% Indels: 0
DB: 18 Gaps: 0

US-10-071-838-2 (1-549) x US-10-839-882-31 (1-1993)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 42 ATGACGCTGTAGAGGTCCGGGCACTGTGGGCAAGAGCGAGAGCATCATATG 101
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 102 AAATACGAAAGGAGACACGAGCTGGCTGCCAGAGCAAGGGCCCTTAACCTTTTGA 161
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 162 AGCTACAAACAACGTCATCTATTGGGATGTATGATGAGACGAGCTGCTCTCTG 221
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 222 ACTGCGCGGAGGAGGAGCAATTCGGCGGAGATCAGCCGAAGAGCAAGTGGTGAT 281
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 282 ATGCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATAGATGAGCGTACAAG 341
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 342 GGAATGCCCATGAACATCCGGGGGCCGATGTGTCACTCTCTGAACACTGAGGAATG 401
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 402 AAGTTGAAAAACCCCGAAGATACAGATCATGAGAGAGAGGCAAGAGTCACTTCAAG 461
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 462 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAGCATATATTCTTCAAG 521
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 522 GATCGATACGGAACCAAGACGCGGAACTACTCCATCTCTCTGCAATGAGAGTAT 581
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
DB 582 AACCCGAGGTGGGCTACTGACAGGAGCTGAGCCACATCGCCGCTGTCTCTCTAT 641
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

DB 642 CTTCCTGAGAGAGATGATCTGCGCACTGGTGACGCTGCTGGCCAGTGAGAGCACTCC 701
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGly 240
DB 702 CTGCAGGATTTACAGCCCAATGGCGGAGCTCCAGGGCTCCAGAACCAACAGAGAG 761
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 762 CATGTGTAGCCACGTCACCAACCAAGACCATGGGGCATCAGAGCAAGAAAGATCATGT 821
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 822 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACCGGATCTCTCGGG 881
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300
DB 882 CTCACCTGCGCTGTGGAGCGTATCTGTAGAAAGCGAAGCAGGCGTTGATGCCGATA 941
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 942 ACAGAATCGCCTTAAAGTTACAGCAGAGCGCTCAGAAAGACGTCACAGTGTGCCCG 1001
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1002 TGGGCACTTTTTCACACCGGTTCTGTGATCCTGGGCCAGGAGTACACTGTGCTC 1061
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1062 AAGCATCTTAGGCGCTTATGAAAGAACTAACAGAAAGAGGAGGAGCTGCCACCCCA 1121
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1122 GCCAAACCCGAGCAAGGCTGTCGATCAGGCTGTGCGGCTTCACTGCGGGAG 1181
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1182 ACCCTGTCAAGGGGAGACAGGAGGCCCTCCAGGCCACAGCCCGGTTCCCGGGCC 1241
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1242 ATTTGTACGCTTCCCGCCACGCGGACCTGTCTTCCACACCTCTCTGTTGGGCT 1301
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1302 GTCCGGGAAGACACTTACCTGTGGGCACTAGGGGTGCCCCAGCCGCTGCTCAG 1361
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1362 GAGGACCTCAGGGTCTCTGAGATTCTGAGTGAATCTCATGCCCGCTCCCAACG 1421
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1422 GACCTGACGTAGAGGCGCTTGTGTTCCGCCATTATGATTTCAGACAGAGCTGGGTG 1481
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
DB 1482 CGTGCCATATCCAGAGAGACCAAGCTGCGCCCTGCTGCAAGCTGAACACCTGCGAG 1541
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1542 CCGGTAGATCGGCTTTCCTGCTGCAACCAAGCATGATTCCGACCAAGGCAACCCCTTCA 1601
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLysGlyLysLeuHisLeu 540
DB 1602 GCTAGGAGCAAGACCGGTGTGCTCCACCTCAGGGCGCTTGCCTTCCAGCTTG 1661
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 1662 GAAAGTTCTAGTTCCTCCAGGCTTC 1688
RESULT 3
US-10-071-838-3
; Sequence 3, Application US/10071838

```

; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2146
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
; NAME/KEY: CDS
; LOCATION: (1)..(1832)
US-10-071-838-3
```

```

Alignment Scores:
Pred. No.: 1.14e-258
Score: 2907.50
Percent Similarity: 89.67%
Best Local Similarity: 89.67%
Query Match: 97.47%
DB: 14
Length: 2146
Matches: 547
Conservative: 0
Mismatch: 2
Indels: 62
Gaps: 1
```

US-10-071-838-2 (1-549) x US-10-071-838-3 (1-2146)

```

QY 1 MetaspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGACGTGGTGAAGGTCGGGAGTGGTGGGACAGAGCGAGACATCATTTATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAGGAGACCGAGCTGGCTGCCAGAGACAGAGGGCCCTTAAGCCTTTTGA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHis----- 53
Db 121 AGCTACACACACACACGTCGATCATTTGGGATTTGTACAGTCCCTCGCCCTCTGGAGTCA 180
QY 53 ----- 53
Db 181 GCCCCACAGAGAGCCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 53 ----- 53
Db 241 CTGAAAGGACAGAGCCAGTCTTTCTGGGGTCCGACCCAGCGCTGGGGCCCTCCAG 300
QY 54 ----- 54
Db 301 GCCCCGTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 60 LeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79
Db 361 CTGACTGCGCGGAGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGACCAAGTGGGT- 419
QY 80 AspMetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr 99
Db 420 GATATGCTGGAGACTGGAGAAATTCACAAAGCAGACAGAAAGCTCATATGATCGAGCGTAC 479
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlu 119
Db 480 AAGGGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAAACATTGAGGAA 539
QY 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysArgSerSer 139
```

```

Db 540 ATGAAGTTGAAGAAACCCCGAAGATACCATATGAAGAGAGGAGAGAGTCTATCT 599
QY 140 GluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159
Db 600 GAGCACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGAAACATATATCTTTC 659
QY 160 ArgAspArgTyrGlyThrLysGlnArgGluLeuHisIleLeuLeuAlaTyrGluGlu 179
Db 660 AGGATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGCAATGAGGAG 719
QY 180 TyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeu 199
Db 720 TATAACCCGAGGTGGGCTACTGCAGGACCTGAGCCACATCGCCCTGTCTCTCTC 779
QY 200 TyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219
Db 780 TATCTTCTGAGAGATGATCATTTCTGGCACTGTGCACTGTGCACTGTGCACTGTGCACTGT 239
QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 259
Db 840 TCCCTGACAGGATTTCAACAGCCCAATGCGGACCGTCCAGGCGGCTCCAAGACCAAG 299
QY 240 GluHisValAlaAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 319
Db 900 GAGCATGTGTAGCCACGTCACCAACCAAGACCATGGGCATCAGACAAAGAAATCTA 339
QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu 359
Db 960 TGTGGCAGTGTTCCTCCCTTAGGCTGCTCATCCGATATGATTTGACCGGATCTCTCTC 379
QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetPro 399
Db 1020 GGGCTCAACCTGCGCTGTGGGACGTGTATCTGTGAAAGCGAAGCGCTGTGATGCCG 419
QY 300 IleThrArgIleAlaPheLysValGlnLysArgLeuThrLysThrSerArgCysGly 439
Db 1080 ATAAACAAGATCGCTTTAAGTTACAGACAGAGCGCTTCAAGAAAGCTCCAGGTGTGGC 459
QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 479
Db 1140 CCGTGGCAGCTTTTGGCAACCGGTTGTTGATACCTGGCCAGGATGAGACACATGTG 499
QY 340 LeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 519
Db 1200 CTCAAGCATCTTAGGCTCTATGAAGAACTTAACAAGAAAGCAGGAGCTGCCACCC 539
QY 360 ProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGly 559
Db 1260 CCAGCCAAACCCGAGCAGAGGTGTCGACATCCAGGCTGTGCGGCTTCACTGCGGG 599
QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 619
Db 1320 AAGACCTCTGCAAGGGGAGACAGGAGCCCTCCAGGCCCAAGCCCGGTCCCGCGG 659
QY 400 ProIleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGly 679
Db 1380 CCCATTGTTGTCAGCTTCCCGCCAGCGGCACTCTCTTCCACACCTGTCTGTGGG 719
QY 420 AlaValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAla 739
Db 1440 GCTGTCGGGAAACACCTACCTGTGGGCACTCAGGCTGTGCCAGCCCGGCTGTGGCT 779
QY 440 GlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro 799
Db 1500 CAGGAGGACCTCAGGTTCTGAGATTCCTGACGTGAACTTCATGCCCGCTCCCA 839
QY 460 ThrAspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 859
Db 1560 ACGGACTGGAAGTAGAGGCGCTTGTTCGCGCATTAATTTCAAGACAGAGCTGTGG 899
QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAla 919
```


Db 1620 GTCCGTGCATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGCG 1679
Qy 500 GGuArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPhe 519
Db 1680 GAGCGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCCAGGGCACCCCTTC 1739
Qy 520 ArgAlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCybGlyLeuHis 539
Db 1740 AGAGCTAGGGACGAACAGCCGTGTGCTCCCACTTCAGGGCCTTGCCCTGCGGCTCCAC 1799
Qy 540 LeuGluSerSerGlnPheProGlyPhe 549
Db 1800 TTGGAAAGTCTCAGTTCCTCCAGGCTTC 1829

RESULT 4
US-10-302-172-692
; Sequence 692, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dimanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich protein-like Nucleic Acids an
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803_1CNCp
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 692
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(1755)
US-10-302-172-692

Alignment Scores:
Pred. No.: 2.25e-253 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.36% Conservative: 1
Best Local Similarity: 96.17% Mismatches: 2
Query Match: 95.54% Indels: 18
DB: 16 Gaps: 1

US-10-071-838-2 (1-549) x US-10-302-172-692 (1-2084)

Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGCTGTAGAGGTGCGGGCAGTGTGGGCAAGAAGCGAGGACATCATTTATG 222
Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCCTTAAGCCTTTTGA 282
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACAAACAACAGCTCATCTATTGGGATTGTACATGAGACGAGCTGCTCTCTTG 342
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGCGGAAGCAATTCGCGGAGATCAAGCCGAAAAGCAAGTGGGTGAT 402
Qy 81 MetLeuGluAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGAGAAATACAAAAGCAGACGAAAAGCTCATAGATGAGCGTACAAAG 462

Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAAATCCGGGGCCCGATGTGTAGTCTCTCTGAACAAGTGAAGAAATG 522
Qy 121 LysLeuLysAsnProGluArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAACATATATTCTTCAGG 588
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLysTyr 180
Db 589 GATCGATACGGAACCAAGCAGCGGAGACTACTCCACATCTCTGGCATATGAGAGATAT 648
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGAGATGATCTTGTGGCACTGTGACGCTGTGCGCAGTAGAGGCACTCC 768
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGCAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTTCCAAGACCAACAGAGAG 828
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGGTAGCCAGTCAACAACCAAGACATGGGGCATCAGGACAAAGAAATCTATGT 888
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTTCCTGTTAGGCTGAGCTGCTCATCCGATATTGATGACGGGATCTCTCGGG 948
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCTGCGGCTGTGGAGCGTGTATCTGTAGAAAGCGGAACAGCGCTTGATGCCGATA 1008
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGAATCGCCTTAAAGTTACAGCAGAACCGCTCAGCAAGACCTCAGGTGGCCCG 1068
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGCACGTTTGTGCAACCGGTTCTGTGATACCTGGCCAGGAGTGAAGACACTGTGCTC 1128
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGGCTCTATGAGAACTTAACAAGAAAGCAGGGGAGCTGCACACCCCA 1188
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1189 GCCAAACCCGAGCAAGGGTGTGTCGATCCAGGCCCTGTGCGGCTTCACTGGCGGAGAG 1248
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGCCACAGCCGGTTCCGCGGCCCC 1308
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1309 ATTTGGTCACTTCCCGCCACAGGCACTGCTTCTTCCACACCTGTCTGTGGGGCT 1368
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACACTAACCTGTGGGCACTCAAGGTGTGCCAGCCCGGCTGCTCAG 1428
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GAGAGACCTCAGGGTCTCTGAGATTCTGTGAGTGAAGTCAATGCCCCCGCTCCCAAGC 1488

QY 461 AspleuaspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGGAAGTAGAGGCGCCCTGTGCTCCGCATATGATTTGACAGAGCTGCTGGGTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1549 CGTGCCATATCCAGAGAGACGACGCTGCCCCCTGCTGGCAGGCTGAACACCCCTCGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGTGAGATCGGCTTTCGTGTCACCCAGCACTGATTCGACCAAGGCGCACCCCTTCAGA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 1669 GCTAGGACGAACAGCCCTGTGCTCCACCTCAGGCGCTTGCTGCGGCTTCCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1729 GAAAGTTCTCAGTCCCTCCAGGCTTC 1755

RESULT 5

US-10-302-172-691
; Sequence 691, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
; FILE REFERENCE: 803 1CNC P
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 691
; LENGTH: 2072
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163) .. (1743)
US-10-302-172-691

Alignment Scores:
Pred. No.: 2.98e-251
Score: 2827.00
Percent Similarity: 95.63%
Best Local Similarity: 95.63%
Query Match: 94.77%
DB: 16
Length: 2072
Matches: 525
Conservative: 0
Mismatches: 2
Indels: 22
Gaps: 1

US-10-071-838-2 (1-549) x US-10-302-172-691 (1-2072)
QY 1 MetaspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGTGTAGAGCTCGCGGCACTTGTGGGCAAGAAGCGAGGACATCATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAGAGGCGCTTAAGCCTTTTGA 282
QY 41 SerTyrAspAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 283 AGCTACAACAACAAGCTCGATCATTTGGGGATTGTACATGACGAGCTGCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 343 ACTGCGCGGAGCGCAAGCAATTCCGCGGAGATCAGCCGAAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAAAAGCTCATAGATCGAGGCTACAA 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db 463 GGAATGCCCATGAACATCCGCGGCGCCGATGTGTCAGTCTCTGTAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGGAGAAAGGCAAGGTCATCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 583 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAGGAAGCATATATCTTCAGG 642
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 643 GATCGATACGGAACCAAGCAGCGGGAATCACTCCACATCTCTGTCATATAGAGATAT 702
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 703 AACCCGAGGTGGCTACTGTCAGGAGCCTGAGCCACATCGCCGCTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 763 CTTCCTGAGGAGATGATCTTGTGGCACTGTGCACTGCTGCGCCAGTGAAGGACACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 823 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGCTCAAGACCAACCAAGAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 883 CATGTGTAGCCACGTCACCAACCAAGACCAATGGGCATCAG----- 924
QY 261 GlyLysCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 937 CTCACCTGCGCTGTGGAGCTGTATCTGTAGAGCGCAAGAGCGTCCAGGTGTGCCCG 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGAATCGCCTTTAAGTTCAAGCAGAGCGCTCAAGAGCGTCCAGGTGTGCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGCAAGTTTGTCAACCGGTTGTTGATACCTGGGCCAGGAGTAGAGCACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1117 AAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAAGCAGGGGACTGCAACCCCA 1176
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
Db 1177 GCCAAACCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCAAGTGGCGGAAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTCTGCAAGGGGACAGGAGCGCCCTCCAGGCCCAAGCCCGGCTTCCCGCGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTGTGTCAGCTTCCCGCAAGGCGACCTGTTCTTCCACACCTGTCTGTGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

Db	1357	GTCCGGGAAGACACCTACCTCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAG	1416
QY	441	GLYGLYPROGLINGLYSERTRPARGPHELEUINTRPANSERMETPROARGLEUPROTHR	460
Db	1417	GGAGGACCTCAGGGTTCTGGAGATTCTCTGACGTGAATCTCATGCCCGCTCCACAG	1476
QY	461	ASPLEUASPVALGLUGLYPROTRPHEARGHISTYRASPBPHEARGLINSERCYESTRPPAL	480
Db	1477	GACCTGGACGTAGAGGGCCCTTGCTTCCGCCATTATGATTTCCAGACAGAGCTGCTGGTTC	1536
QY	481	ARGALAIIESERGLINLUASPGLINLEUALAPROCYSTRPGLNALAGLUHISPROALAGLU	500
Db	1537	CGTGCCATATCCCAAGAGACCAGCTGGCCCCCTGTGCGCAGGCTGAACACACCTGCCGAG	1596
QY	501	ARGVALARGSERALAPHEALAPROSETRHASPSERAASPGINGLYTHRPROBHEARG	520
Db	1597	CGGGTGAGATCGGCTTTGCTGTCACCCAGCACTGATTCGACACAGGACACCCCTTCAGA	1656
QY	521	ALAARGASPGIUGLINPROCYSAALAPROTHRSEGLYPROCYSEUCYSGLYLEUHI	540
Db	1657	GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCTTGCCCTCTGCGGCTTCACCTG	1716
QY	541	GLUSERSERGLINPHEPROGLYPHE	549
Db	1717	GAAGTTCCTCAGTTCCTCCAGGCTTC	1743

```

RESULT 6
US-10-108-260A-676
; Sequence 676, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-676

```

Alignment Scores:	
Pred. No.:	4.01e-251
Score:	2827.00
Percent Similarity:	95.63%
Best Local Similarity:	95.63%
Query Match:	94.77%
DB:	16
Length:	2647
Matches:	525
Conservative:	0
Mismatches:	2
Indels:	22
Gaps:	1

US-10-071-838-2 (1-549) x US-10-108-260A-676 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 755 ATGACGTGTAGAGGTCCGGGCGAGTGTGTGGCAACAAGCGAGAGACATCATTTATG 814
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProlLysProPheArg 40
DB 815 AAATACGAAAGGACACCAGCTGGGCTGCCAAGAGACAAGGGCCCTAAGCCTTTTCGA 874
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProlLeu 60
DB 875 AGCTACACACACACGTCGATCATTTGGGGATTGTACATGAGACCGAGCTGCTCTCTGTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 935 ACTGCGGGAGGCGAACAATTCCGGCGGAGATCACCGAAAAGCAAGTGGGTGAT 994
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 995 ATGCTGGAGACTGGGAGAATAACAAAGACAGCGAAAGCTCATAGATCGAGCGTACAG 1054

QY	101	GIYMeCPromETAsnIleArgGlyPROMETTrpSerValLeuLeuAsnIleGluGluMet	120
Db	1055	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTCGAACAACATGAGAAATG	1114
QY	121	LYbLeuLYAsnProGlyArgTyrGlnIleMetLYGluLYbLYLYbArgSerSerGlu	140
Db	1115	AAGTTGAAAAACCCCGGAAGATACCAGATCATGAAGAGAAGGCAAGAGTCACTCGAG	1174
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLYbHisIlePhePheArg	160
Db	1175	CACATCCAGCCGATCGACCCGGACGTAAAGCCGGACATTAAAGAACATATATTCTTCAGG	1234
QY	161	AspArgTyrGlyThrLYSGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	1235	GATCGATACGGAACCAAGCAGCCGGAACCTACCATCTCTCGCATATGAGAGTAT	1294
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200
Db	1295	AACCCGGAGGTGGGCTACTGCAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT	1354
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	1355	CTTCCTGAGAGGATGCATTCTGGCACTGGTGCACCTGCGCAGTGAGAGCACTCC	1414
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	1415	CTGCAGGATTTCACAGCCCAATGCGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG	1474
QY	241	HisValValAlaThrSerGlnProLYSThrMetGlyHisGlnAspLYbLYbAspLeuCys	260
Db	1475	CATGTGTAGCACGTCACAACCCAGACCATGGGGCATCAG-----	1516
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	1517	-----ATCTCTCTCGGG	1528
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle	300
Db	1529	CTCACCCCTGCCCTGTGGGACGTGTATCTGGTAGAAGCGGAACAGCGCTGTGATCCGATA	1588
QY	301	ThrArgIleAlaPheLYbValGlnGlnLYbAspLeuThrLYSThrSerArgCysGlyPro	320
Db	1589	ACAAGAATCGCCTTAAAGTTCCAGCAGAAAGCGCTTCAAGAAGCTCCAGGTGTGCCCG	1648
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1649	TGGGCACGTTTTCGCAACCGGTTCTGTGATACCTGGGCCAAGGATGAGGACACTGTGCTC	1708
QY	341	LYbHisLeuArgAlaSerMetLYbLYbLeuThrArgLYSGlnGlyAspLeuProProPro	360
Db	1709	AAGCATCTTAGGGCTCTATGAAGAACTAAAGAAAGCAGGGGGAACCTGCACACCCCA	1768
QY	361	AlaLYbProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLYbLYs	380
Db	1769	GCCAAACCCGAGCAGAGGTCTCGGCATCCAGGCTGTGCCGCTTACAGTGTGGGGAAG	1828
QY	381	ThrLeuCysLYSGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1829	ACCCTCTGCAAGGGGACAGGCAAGCCCTCCAGGCCACACAGCCCGGTTCCCGCGGCC	1888
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1889	ATTTGTACGCTTCCCGCCACCGGGCACTCTCTTCCACACACCTGTCTCTGTGGGGCT	1948
QY	421	ValArgGluAspThrTyrTrpValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1949	GTCGGGAAGACACTACCTGTGTGGCACTCAGGAGTGTGCCAGCCCGGCTGTGCTCAG	2008
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	2009	GGAGGACCTCAGGGTTCTGTGAGATTCTGTGCACTGGAACCTCCATGCCCCGCTCCCAACG	2068

QY 461 AspleuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 2069 GACCTGGACGTAGAGGCGCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 2129 CGTGCCATATCCACGAGGAGCACAGCTGGCCCCCTGTGCGCAGGCTGAACACCCCTGCGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyTrpPheArg 520
Db 2189 CCGGTGAGATCGGCTTCGCTGCACCCAGCACTGATTCGACACAGGGCACCCCTTCAGA 2248
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 2249 GCTAGGAGCAGACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335

RESULT 7
US-10-094-466-61

/ Sequence 61, Application US/10094466
/ Publication No. US20030203363A1
/ GENERAL INFORMATION:
/ APPLICANT: Spytek et al.
/ TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
/ TITLE OF INVENTION: AND METHODS OF USING
/ FILE REFERENCE: 21402-290D
/ CURRENT APPLICATION NUMBER: US/10/094,466
/ PRIOR FILING DATE: 2002-03-07
/ PRIOR APPLICATION NUMBER: 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/288,148
/ PRIOR FILING DATE: 2001-05-02
/ PRIOR APPLICATION NUMBER: 60/274,849
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: 60/275,235
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: 60/338,375
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/275,579
/ PRIOR FILING DATE: 2001-03-13
/ PRIOR APPLICATION NUMBER: 60/335,302
/ PRIOR FILING DATE: 2001-10-31
/ PRIOR APPLICATION NUMBER: 60/275,601
/ PRIOR FILING DATE: 2001-03-13
/ PRIOR APPLICATION NUMBER: 60/276,000
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/277,338
/ PRIOR FILING DATE: 2001-03-20
/ Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: PatIn 2.1
/ SEQ ID NO 61
/ LENGTH: 1752
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (31) .. (1678)
US-10-094-466-61

Alignment Scores:

Pred. No.: 2.29e-247 Length: 1752
Score: 2784.00 Matches: 515
Percent Similarity: 96.17% Conservative: 13
Best Local Similarity: 93.81% Mismatches: 21
Query Match: 93.33% Indels: 0
DB: 15 Gaps: 0

US-10-071-838-2 (1-549) x US-10-094-466-61 (1-1752)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 31 ATGACGCTGTAGAGGTCCGCGGTAGTTGGTGGGCAACAAGCGAGACATCATATATG 90
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 91 AAATACGAAAAGGACACCCGAGCTGGCTGCCAGAGACAAGGGCCCTTAAGTCTTTGGA 150
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 151 AGCTACACACACACGCTGATCATTTGGGATTTGTACAGAGACGAGCTGCTCCTCTG 210
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 211 ACTGCGCGGAGGTGAAGCAATTCCGCGGAGATCAGCCGAAAGACAAAGTGGGTGAAA 270
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 271 ATGCTGGAGATGGACACCTACAAAACAGACAAAGCTCATAGATCGAGCGTACCAG 330
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db 331 GGAATTCCTCCATGAACATCCGGGCGCCGATGTGTCAGTCTCCTGAACATGAGGAATC 390
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 391 AAGTTGAAAACCCCGAAGATACAGATCATGAAGAGGACAAAGGATCATCTGAA 450
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 451 CACATCCAGCAGATGACCTCGACGTAAGCGGACATTAAAGAGCATATATCTTCAGG 510
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 511 GATCGATACGAAACCAAGCAGCGGAACTTACTTACATCTCTGGCGGTATGAGAGTAT 570
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
Db 571 AACCCGAGGTGGGCTACTGACGGAGCTGAGCCACATCGCCGCTGTCTCTTAT 630
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 631 CTTCCTGAGGAGATGATCTTGCGCACTGTGACAGCTGCGCCAGTGAGGCACTCC 690
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 691 CTGCAAGGTTTTCACAGCCCAATGCGGAGCCGTCCAGGGGCTCCACAGCAACAGAG 750
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 751 CATGTGTAGCCAGCTCACTACCAACCAACCATGTGGCATCAGACCAAGAAATCTATGT 810
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 811 GGGCAGTGTTCCTTAAAGCTTCAGTACGCGCTCATCCGATATTGATGACGGATCTCTCGG 870
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 871 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAGCGAACAGCGCTTGATGCCGATA 930
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 931 ACAAGAAATCGCCTTAAAGTTCACTAGAGCGCTCACGAAGACGTCCAGGTGTGCGCCG 990
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 991 TGGGCACTTTTGAACCGGTTGTTATGCTGCGCCAGGAGTATGACACTGTGCTC 1050
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1051 AAGCATCTTAGGCGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCAACCCCA 1110

QY 361 AlAlysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1111 GCCAAACCCGAGCAGAGGTCGTGGCATCCAGGCTGTGCCAGCTTCACGTGGCGGAAG 1170
QY 381 ThrLeuCySylsGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1171 ACCCTCTGCAAGGGGACAGCAGCAGCCCTCCAGGCCACCAAGCCGGTTCATGGGCC 1230
QY 401 IleTrpSerAlaSerProProAlaArgAlaProArgSerSerThrProCySProGlyGlyAla 420
Db 1231 ATTTGGTCAAGCTTCCCCGCCACCGGGCACCTGTTCTTCCACACCTCTGCTGGTGGGCT 1290
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1291 GTCCGGAGACACCTACCTCTGTGGCACTCAAGGTGTGCCAGCCCGCTGTGGCTCAG 1350
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1351 GGAGGACCTCAGGGTTCCTGGAGATTCCTGCACTGGAATCCATGCCCCCTCCCAACG 1410
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCySTrpVal 480
Db 1411 GACCTGACGTAGGGGACCCCTTGTTCCGCCGTATGATTTCAGACAGAGCTGTGGGTC 1470
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCySTrpGlnAlaGlnHisProAlaGln 500
Db 1471 CGTGCCATATCCACGAGAGACCAAGCCGGCCACTGTGGCAGGCTGAACACCTGTGGAG 1530
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1531 CGGGTAGATCGGCTTTTCAGTGCAGCCCGCAGCACTGATTCGACACAGGCGACCCCTTCAGA 1590
QY 521 AlaArgAspGluGlnProCySAlaProThrSerGlyProCySLeuCySgLyLeuHisLeu 540
Db 1591 GCTAGGAGCAACAGCAGTGTGCTCCCACTCAGACCTTGCCCTGCGGCTCCACTTG 1650
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1651 GAAAGTCTCAGTTCCTCCACAGGCTTC 1677

RESULT 8
US-10-071-838-5
; Sequence 5, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1538)
US-10-071-838-5

Alignment Scores:
Pred. No.: 2.56e-246 Length: 1862
Score: 2773.00 Matches: 515
Percent Similarity: 93.81% Conservative: 0

Best Local Similarity: 93.81% Mismatches: 0
Query Match: 92.96% Indels: 34
DB: 14 Gaps: 1
US-10-071-838-2 (1-549) x US-10-071-838-5 (1-1862)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGAGCGTGTAGAGTCGCGGCGCAGTTGTGGCACAAGAGCAGAGCATCATATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGACAGGGGCCCTTAAGCCTTTTCGA 120
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 121 AGCTACAAACAACGTCGATCATTTGGGATTGTACATGAGACGGAAGCTGCTCTCTTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTGCGCGGAGCGAAGCAAAATTCCGCGGAGATCAGCCGGAAGACAAAGTGGTGAT 240
QY 81 MetLeuGluAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAGAAAGCTGATAGAGCTAGCAAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 301 GGAATGCCCATGAAATATCCGGGGCCCGATGTGTCACTCTCTGAACATTTGAGGAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAAACCCCGAAGATACCATGATGAAGGAGAGGCAAGGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATGACCGGAGCAGGACGGAACATTGAAGGAGAGCATATATTCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGATAT 540
QY 181 AsnProGluValGlyTyrCySArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 541 AACCCGAGGTGGCTACTGACGGGACCTGAGCCACATCGCCGCTTCTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCCTGAGGAGATGATCTTCTGGGCACTGGTGACGTGCTGGCCAGTGAGAGGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 661 CTGACGAGATTTCACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAAGCAACACAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCyS 260
Db 721 CATGTGTAAGCCAGTCACAAACCAAGCATGGGGCATCAG----- 762
QY 261 GlyGlnCySserProLeuGlyCySLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 762 ----- 762
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 763 -----TATCTGTAGAAGCGCAACAGCGCTGTATGCCGATA 798
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCySgLyPro 320
Db 799 ACAAGATCGCTTTAAGTTACAGAGAGCGCTCACGAAGACGTCACAGTGTGGCCCG 858
QY 321 TrpAlaArgPheCySAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 859 TGGGACAGTTTTCACACCGGTTCTGTGATACCTGGGCCAGGAGATGAGACACTGTGCTC 918

QY 341 LysHisLeuArgAlaSerNectIlySlyLeuThrArgLySGInGIyAspLeuProProPro 360
Db 919 AAGCATCTTAGGGCTCTATGAAAGAACTAACAGAAAGCAGGGGAGCTGCCACCCCA 978
QY 361 AlAlysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgLyGlyLys 380
Db 979 GCCAAACCCGAGCAGAGGTCGTGGCATCCAGGCTGTGCCGCTTCCAGTGGCGGGAAG 1038
QY 381 ThrLeuCySlySGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1039 ACCCTCTGCAAGGGGGACAGGAGGCCCTCCAGGCCCAACGACCCGGTTCCCGCGGCC 1098
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1099 ATTGGTCAGCTTCCCGCCACCGGCACTCTGTTCTTCCACACCTGTCTGGTGGGCT 1158
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1159 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCGCGCTGTGGCTCAG 1218
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1219 GGAGGACCTCAGGGGTCTCTGAGATTCTCTGACGTGAACTCCATGCCCGCTCCCAAG 1278
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1279 GACCTGACCTAGAGGGCCCTTGGTTCGCCATTATGATTTCACAGACGCTGTGGTGC 1338
QY 481 ArgAlaIleSerGlnGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1339 CGTGCCATATCCAGAGGAGACAGCTGGCCCCCTGCTGGCAAGGCTGAACACCTCGGAG 1398
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1399 CGGGTGAGATCGGCTTTCCTGCTGACCCAGCAGACTGATTCGACAGGGCACCCCTCAGA 1458
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 1459 GCTAGGAGCAGAACAGCCGTGTCTCCACCTCAGGGGCTGTGCTCTGCGGCTTCACCTG 1518
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1519 GAAAGTTCAGTTCCTCCACAGGCTTC 1545

RESULT 9

US-09-962-436-562
; Sequence 562, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR APPLICATION NUMBER: 2001-09-25
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 562
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-562

Alignment Scores:

Pred. No.: 7,74e-192
Score: 2191.50
Percent Similarity: 88.20%
Best Local Similarity: 81.60%

Length: 7878
Matches: 408
Conservative: 33
Mismatches: 58

Query Match: 73.47% Indels: 1
DB: 9 Gaps: 1
US-10-071-838-2 (1-549) x US-09-962-436-562 (1-7878)
QY 1 MetaSpValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGACATGCTGTAGAGATGACAGATTGTTGACAGGACAGAGCGGAAGACATATTATG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCAAGCTGGCTGCCAGAGGACAAAGGGGCTGAGCCCGTT--- 1813
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTGTCATGAGACGACCTGCTCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValasp 80
Db 1874 ACTGCAAGGAGGCGAAGAAATTCGGCGGAGATGACACGAAGCAAGCAAGTGATGCAA 1933
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGCTGGAGAGATGGAGACATATAGCACAGTACCAAACTCATGATCGAGTCTACAA 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 1994 GGAATTCATGAACATCCGGGGCCCGGTGTGTCAGTCTCTGAACTTCAGGAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGAAGATACAGATCATGAGAGAGAGGGGCAAGAGTCACTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCAACATCGACCTGAGCTGAGAGCACTCTCCGAAACCATGCTTCTTAGG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATCGATATGAGCCAAAGCAGAGGAACTATTCTACATCTCTGCTTATTCGAGATAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 2234 AACCCGAGGTGGCTACTGACGGACCTGAGCCACATCACCGCTTCTCTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCTGAGAGAGACGATTCGTGGCACTGTGTGCACTGTGGCCAGTGAGAGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGATTCACAGCCCAATGTGGGACAGTCCAGGGGCTCCAAAGCAACAGAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 2414 CATGTGTAACCAAGTCAACCAACCAACCATGTGGCATTCAGACAAAGAAAGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTGTAGGCTGCTCTCCGAACCTGATTGACGGGATCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 2534 CTCACCTGCGCTGTGGACGTGATTGTGGAAGAGAAACAGGTGTGATGCCATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATGCTCTTAAGTTACAGCAGACGCTCATGAACATTCAGGTGTGGCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGGCAAGTCTGCGAACAATTCCTGATACCTGGCCATGAACGATGACACCGTGTCT 2713

QY	341	LYSHISLEUARGALASERMETLYSLYSLEUTHRRARGLYSGINGLYASPLeuPROProPro	360
DB	2714	AAGCATCTTAGGGCCCTTACGAGAAGAACTAACAGGAAGCAAGGGGACCTGCCACCCCCA	2773
QY	361	ALALysPROGLINGLYSERSERALASERARGPROVALPROALASERAGLYGLYLS	380
DB	2774	GCCAAACGGAGCAAGGGTCTTGGCACCCAGGCGCTGTGCGGCTTCACGTGTGGGAG	2833
QY	381	THRLeuCYLSYSGLYASPARGINALAPROPROGLYPROPROALARGPHEPROARGPRO	400
DB	2834	ACCCTGTCAAGGGGTATAGCAGGCCCTCCAGGCCCAACAGTTCAGCGGCC	2893
QY	401	IleTRPserALASERPROPROARGALAPROARGSERTHRPROCYSPROGLYGLYALA	420
DB	2894	ATTGTCTAGCTTCCCGCCATGGGCATCTCGTTTTTCCAGCCCTGTCTGTGGGCT	2953
QY	421	VALARGGLUASPThrTYRPROVALGLYTHRGLNGLYVALPROSERPROALALEUALAGLN	440
DB	2954	GTCCGGGAAGACACGTACCTGTGGGCACTCAGGGTGTGCCAGCTGGCCTGGCTCAG	3013
QY	441	GLYGLYPROGLINGLYSERTRPARGPHELEUGLINTYPANSMETPROARGLEUPROTHR	460
DB	3014	GGAGGACCTCAGGGTCTCTGAGATTCCTGTGAGTGAAGTCAATGCCCGGCTCCCAACG	3073
QY	461	ASPLEUASPVALGLUGLYPROTRPHEARGHISTYRASPHEARGGLNSERCYETRYVAL	480
DB	3074	GACCTGATATAGGGGGCCCTTGCTTCCCCCATTTATGATTTTGAACGAGACTGCTGGGTC	3133
QY	481	ARGALAILESERGLINGLYASPGLNLEUALAPROCYETRPGLNALAGLUHISPROALAGLU	500
DB	3134	CGTGCCATATCCAGGAGACCACTGGCCACCTGCTGGCAGGCTGAACACTGCCGAGAG	3193

RESULT 10
US-10-283-975A-272

```

; Sequence 272, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; PRIORITY FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 8201
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-272

```

Alignment Scores:	
Pred. No.:	1.85e-186
Score:	2133.50
Percent Similarity:	77.23%
Best Local Similarity:	71.45%
Query Match:	71.52%
DB:	17
Length:	8201
Matches:	408
Conservative:	33
Mismatches:	58
Indels:	72
Gaps:	2

US-10-071-838-2 (1-549) X US-10-283-975A-272 (1-8201)

[illegible]

Db	1757	AA	GTATGACAAGGACACACCGAGCTGGGCTGCCAGAGACAAGGGCGCTGAGCCGTT---	1813
Qy	41	Ser	TrpAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu	60
Db	1814	GGA	TCACACAGCAGCATTTGATCGTTTGGCATTTTGCATGAGACGGAGCTGCCTCTGTG	1873
Qy	61	Thr	AlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
Db	1874	ACT	CACGGGAGCGGAGAAATTCGGCGGGAGATGACACGAACGAGCAAGTGGAATGAA	1933
Qy	81	Met	LeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
Db	1934	ATG	CTGGGAGATGGGAGACATATAAGCACAGTAGCAAACTCATAGATCGAGTGTACAAG	1993
Qy	101	Gly	MetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet	120
Db	1994	GGA	ATTCCTCCATGAACATCCGGGGCCCGGTGTGTGTCAGTCTCTGAACTTCAGGAATC	2053
Qy	121	Lys	LeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
Db	2054	AAG	TGAAAAACCCGGAGATACCATCATGAGAGAGGGGCAAGAGTCACTGAA	2113
Qy	141	His	IleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	2114	CAC	ATCCACCATCGACTGACCTGGACGTGAGACGACTCCGGAACCATGTCTTTTAGG	2173
Qy	161	Asp	ArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	2174	GAT	CGATATGGAGCCACGACGAGGGAATATTCTACATCTCTCGGCTATTCCGAGTAT	2233
Qy	181	Asn	ProGluValGlyTyrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTyr	200
Db	2234	AAC	CGAGGTGGCTACTGACGGGACTGAGCCACATCACCGCTGTTCCTCTTAT	2293
Qy	201	Leu	ProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	2294	CTG	CTGAGAGAGACGCACTTCTGGGCACTGCTGACAGTCTGGCCAGTGAAGGCACTCC	2353
Qy	221	Leu	GlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	2354	CTG	CCAGATTCACACAGCCCAATGGTGGACAGTCCAGGGGCTCCAAGCAACAGAGAG	2413
Qy	241	His	ValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	2414	CAT	GTGTACCCCAAGTACACCAACCAAGACCATGTGGCATCAGGACAAAGGAAGTCTATGC	2473
Qy	261	Gly	GlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	2474	GGG	AGTGTCTCTGTTAGGCTTCTCTCCGGAACCTGATTGACGGGATCTCTCGGG	2533
Qy	281	Leu	ThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle	300
Db	2534	CTC	ACCCCTGCGCTGTGGAGCTGATTGTGTGAAGAGAACAAGGTGTGATGCCAATA	2593
Qy	301	Thr	ArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	2594	ACC	AGCATTTGCTCTTAAGGTTCAAGACAGCGCCTCATGAAGACATCCAGGTGTGGCTG	2653
Qy	321	Tyr	AlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	2654	TGG	GACGCTCTGGAAACCAATTCTTGATACCTGGGCCATGAACGATGACACCGTGTCTC	2713
Qy	341	Lys	HisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	2714	AAG	CATTTAAGGCTCTACGAAGAACTAACAAGGAAGCAAGGGGACTGCCACCCCA	2773
Qy	360	-----	-----	360
Db	2774	GGC	CAACAGCCCTGGGACGAAGTGTGTGGCAGGAAGCCCCCAGCCAGTCTGAACCTG	2833
Qy	360	-----	-----	360

QY	360	-----	360
Db	2774	GGCCACAGCCCTGGACGAGGTGTGTGGCAGGAGCCCCCAGCCAGTCTGAACCTG	2833
QY	360	-----	360

360 ----- 360

Db 2834 GGGGAGTCCAGAGCCACCACCATGCCCCAACGGCTTCCCATGCGAGCAGACAC 2893
QY 360 -----
Db 2894 ACCCCTCCCTGTGGATCAGACACTACAGCGGTGTCTCAGTGTACAGACACAGGGCC 360
QY 361 -----
Db 2954 ACACAGAGACCCCAAGACTCCAGAGATGACGCCAAAGCGAGCAGAGGTCCTTGACAC 3013
QY 370 rArgProValProAlaSerArgGlyGlyThrLeuCysLysGlyAspArgGlnAlaPr 390
Db 3014 CAGGCTGTGCGGCTTACGCTGTGGAGAACCTCTGCAAGGGGTATAGGAGGCCCC 3073
QY 390 oProGlyProProAlaArgPheProArgProIleTyrSerAlaSerProProAlaPr 410
Db 3074 TCCAGGCCCAACGAGCCCAAGTCCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCATC 3133
QY 410 oArgSerSerThrProCysProGlyGlyAlaValArgGlnAspThrTyrProValGlyTh 430
Db 3134 TCCTTTTCCACGCCCTGTCTGTGGGGCTGTCCGGAGAACAACGTAACCTGTGGGCAC 3193
QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrArgPheLe 450
Db 3194 TCAGGGTGTGCCAGCCTGGCCCTGCTCAGGAGAACCTCAGGGTCTTGAGATTCT 3253
QY 450 uGlnTyrAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTyrPheAr 470
Db 3254 GGAGTGAAGTCAATGCCCCGGCTCCCAACGAGCTGATATAGGGGGCCCTTGTTCCC 3313
QY 470 gHisTyrAspPheArgGlnSerCysTyrValArgAlaIleSerGlnGlnAspGlnLeuAl 490
Db 3314 CCATTATGATTTTGAACGGAGCTGTGGTCCGTCCATATCCAGAGAGACCAAGCTGGC 3373
QY 490 aProCysTyrGlnAlaGlnHisProAlaGlu 500
Db 3374 CACCTGCTGGCAGGCTGAACACTGCGGAGAG 3404
RESULT 11
US-10-098-841-51
; Sequence 51, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dimañac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_fl_genes version 1.0
; SEQ ID NO 51

; LENGTH: 8408
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2859)..(6116)
US-10-098-841-51
Alignment Scores:
Pred. No.: 1.91e-186 Length: 8408
Score: 2133.50 Matches: 408
Percent Similarity: 77.23% Conservative: 33
Best Local Similarity: 71.45% Mismatches: 58
Query Match: 71.52% Indels: 72
DB: 13 Gaps: 2
US-10-071-838-2 (1-549) x US-10-098-841-51 (1-8408)
QY 1 MetAspValValGluValAlaGlySerTyrTrpAlaGlnGluArgGlnAspIleIleMet 20
Db 1697 ATGACATGATGATGAGAATGACAGATAGTTTGACAGGCACAGAGCGGAGACATATATG 1756
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGlnAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAGAGGACACCCAGCTGGCTGCCAGAGACCAAGGGGCTGAGCCCTT--- 1813
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 1814 GAATCAACAGCAGCATTCGTTTGGCATTTTGACATGACAGGAGCTGCTCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTyrValAsp 80
Db 1874 ACTGCAAGGAGGCGAAGAAATTCGGCGGAGATGACACGAGCAGCATGATGAA 1933
QY 81 MetLeuGlyAspTyrGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGCTGGAGAAATGGAGACATATTAAGACAGTAGCAAACTCATATGAGTACAG 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GAATTCCTCCATGAACATCCGGGGCCCGGTGTGTCTCCTCTGAAATTCAGGAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAATAACCCCGAGATACCATGATCATGAAGAGAGGGCAAGAGTCTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACACATCGACTGACGTGAGACGACTCTCCGAAACATGTCCTTAAAG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATGATATGAGCCCAAGCAGAGGAATAATTCTACATCTCTGCTTTCGAGATAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db 2234 AACCCGAGGTGGCTACTGACGGACCTGAGCCACATCACCGCTTGTCTCTTAT 2293
QY 201 LeuProGlnGluAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCTGAGGAGGACGATTCGGGCACTGTGTGACGCTGTGCGCAGTAGAGGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGGATTCACAGCCCAATGTGGGACAGTCCAGGGGCTCCAAAGCAACAGAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 2414 CATGTGTATCCCAAGTCAACCCCAAGCATGTGGCATCAGACCAAGAGAGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
Db 2474 GGGCAGTGTCTCGTTAGGCTGCTTCTCCGAACTGATGACGGAGATCTCTCTGGG 2533

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluAlaLeuMetProIle 300
Db 2534 CTCACCCCTGCGCCTGTGGACGTGTATTGTGTGAAGAGAACAGAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheIleValGlnGlnIleAspLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATGTCTTAAGGTTCAAGAGAGCGCTCATGAAGACATCCAGGTGTGGCCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGGCACGTCTGCGAACCAATTCTTCGATACCTGGGCCCATGAACGATGACACCGTCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATCTTAGGGCCTTAAGAACTAAAGAAAGAAAGAGGAGGAGCTGCCACCCCA 2773
QY 360 ----- 360
Db 2774 GGCCCAACAGCCCTGGGACGAAGGTGTGTGGCAGGAAGCCCAAGCCAGTCTGAACCCCTG 2833
QY 360 ----- 360
Db 2834 GGGGCACTCCAGAGCCACCACCATGCCCCAACGGCTTCCCATGCCAGGACAGCACAC 2893
QY 360 ----- 360
Db 2894 ACCCTCCTCTGGATTCAGACAGACTACAGCGCTGTCTCAGTGTACAGACCAAGGGCC 2953
QY 361 -----AlaLysProGluGlnLysSerSerAla 370
Db 2954 ACACAGAGACCCCAAGACTCCAGAGATGACGCCAAAGCGAGCAAGGCTCTTGACACC 3013
QY 370 LArgProValProAlaSerArgLysGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
Db 3014 CAGGCTGTGCGGCTTCAAGTGTGGAGAACCTCTGCAAGGGGTATAGGACGGCCCC 3073
QY 390 OProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProAlaPr 410
Db 3074 TCCAGGCCCAACAGCCAGTTCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCATC 3133
QY 410 OArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyTh 430
Db 3134 TCGTTTTCACGCGCCTGTCTGTGGGCTGTCCGGAGACACAGTACCCTGTGGGCAC 3193
QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyLysProGlnGlySerTrpArgPheLe 450
Db 3194 TCAGGGTGTGCCAGCTGGCCCTGAGTCAAGGAGGACCTCAGGGTCTCGAGATTCTC 3253
QY 450 uGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheAr 470
Db 3254 GGAAGTGAAGTCAATGCCCCCGGCTCCCAACGACCTGATATAGGGGCCCTTGTTCCC 3313
QY 470 gHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAl 490
Db 3314 CCATTATGATTTTGAACGAGCTGCTGGTCCCTGATATCCAGAGAGACCAAGCTGGC 3373
QY 490 aProCysTrpGlnAlaGluHisProAlaGlu 500
Db 3374 CACCTGTGGCAGGCTGAACACTGCGGAGAG 3404

RESULT 12
US-10-098-841-50
; Sequence 50, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 50
; LENGTH: 8420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2859) .. (6128)
US-10-098-841-50

Alignment Scores:
Pred. No.: 1,91e-186 Length: 8420
Score: 2133.50 Matches: 408
Percent Similarity: 77.23% Conservative: 33
Best Local Similarity: 71.45% Mismatches: 58
Query Match: 71.52% Indels: 72
DB: 13 Gaps: 2

US-10-071-838-2 (1-549) x US-10-098-841-50 (1-8420)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGACATGTGAGAGATGACAGATAGTTTGACAGGCACAGAGCGGAGAGACATATTATG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTGAGCCCGTT-- 1813
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 1814 GGAATCAACAGCAGCATGTATCGTTTGGCATTTTGATGAGAGAGAGCTGCCTCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGACAGGAGGCGAAGAAATTCCGGCGGAGATGACACGACGAGCAAGTGTGATA 1933
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGCTGGAGATGGAGACATATATAGCAGTAGCAAACTCATAGATCGAGTGTACAG 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCATGAAATCCCGGGCGCGGTGTGTGTCAGTCTCTGAAACATTGAGAAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGGCGCAAGGTGATGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCACTGACCTGACGTGAGAGACGACTCTCCGGAACCATGTCTTCTTAGG 2173

QY 161 AsparGlyThrGlyThrIleArgGluLeuHisIleLeuLeuAlaTyrGluTyr 180
Db 2174 GATCGATATGAGCCAGACAGAGGAAGTATCTTACATCTCTGGCCTATTCGAGTAT 180
QY 181 AsnProGluValGlyTyrCysAspLeuSerHisIleAlaAlaLeuPheLeuTyr 2233
Db 2234 AACCCGAGGTGGCTACTGACAGGACCTGACCATCACCTGCTTCTCTCTATTCGAGTAT 2233
QY 201 LeuProGluGluAspAlaPheThrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 2293
Db 2294 CTGCTGAGAGAGAGCATTCGGGACCTGCTGAGCTGCTGGCCAGTGAAGAGGACTCC 2293
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 2353
Db 2354 CTGCCAGGATTCACAGCCCAATGCTGGACAGTCCAGGGCTCCAAAGACCAAGAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLeuPheLeu 2413
Db 2414 CATGTGTACCAAGTACAAACCAAGACCATGTGGCATCAGAGGCTCCAAAGACCAAGAG 2413
QY 261 GlnGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 2473
Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTCTCCGAACTGATTCAGCGGATCTCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuThrAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 2534 CTCACTCTGCGCTGTGGAGCTGATTTGTGGAAGAGACAGGTGTGATGCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATTCCTTAAAGTTCAAGAGAGCGCTCATGAACATCCAGGTGTGGCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 2653
Db 2654 TGGCAGCTGTGCGAACCATTCTGATACCTGGGCCATGAACATGACACCTGTGCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATCTTAGGGCTCTTACGAAGAACTAACAGAGAGGAGGACCTGCCACCCCA 2773
QY 360 -----
Db 2774 GGGCCAAAGCCCTGGAGCAAGGTGTGTGGCAGAGGCCCCAGCCAGTCTGAACCTG 2833
QY 360 -----
Db 2834 GGGCAGTCCAGAGCAACCAATGCCCAACGGCTTCCCATGCCAGGACACAC 2893
QY 360 -----
Db 2894 ACCCTCCCTGTGGATCAGAGACTACAGCGCTGTGTGTCAGTGTGACACAGGGGCC 2953
QY 361 -----
Db 2954 ACACAGAGACCCAGAGACTCCAGATGTCAGCAACGCCGAGCAAGGTCTTGGCAC 3013
QY 370 ArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPro 390
Db 3014 CAGGCTGTGCGGCTTACGTGTGGGAAGACCTCTGCAAGGGGTATAGGAGGGCCCC 3073
QY 390 GProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPro 410
Db 3074 TCAGAGCCCAAGCCAGTTCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCATC 3133
QY 410 OArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTrpProValGlyTh 430
Db 3134 TCGTTTTCACGCTGCTGTGGGCTGTCCGGGAAGACACGTAACCTGTGGCAC 3193
QY 430 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLe 450
Db 3194 TCAGGCTGTGCCAGCTGTGCTGCTCAGGAGGACCTCAGGGTCTCTGAGATTCCT 3253
QY 450 uGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheAr 470

Db 3254 GGAGTGAAAGTCAATGCCCGGCTCCCAACGAGCCTGATATAGGGGCCCTTGTTC 3313
QY 470 GHisTrpAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAl 490
Db 3314 CCATTATGATTTGAACGAGAGCTGCTGGTCCGTCATATCCAGAGACCACTGCG 3373
QY 490 AProCysTrpGlnAlaGluHisProAlaGlu 500
Db 3374 CACTGTGCAGGCTGAACACTGCGAGAG 3404

RESULT 13
US-09-764-869-1824
/ Sequence 1824, Application US/09764869
/ Patent No. US20020061521A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC007
/ CURRENT APPLICATION NUMBER: US/09/764, 869
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 2442
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1824
/ LENGTH: 9805
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-764-869-1824

Alignment Scores:
Pred. No.: 1.58e-128 Length: 9805
Score: 1507.50 Matches: 487
Percent Similarity: 20.97% Conservative: 6
Best Local Similarity: 20.71% Mismatches: 17
Query Match: 50.54% Indels: 1844
DB: 9 Gaps: 11

US-10-071-838-2 (1-549) x US-09-764-869-1824 (1-9805)

QY 24 LysGlyHisArgAlaGlyLeuProGluAspLysGly-----ProLysProPheArgSer 41
Db 2492 CAGGGCTCCAGAGCCCAAGACCAGCATCCGGCGGCTCTGGAAAGCTTGACAGCTCC 2551
QY 42 TyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeuThr 61
Db 2552 GCTAACTCAACATGCTCTATTG-----
QY 62 AlaArgGluAlaLys-GlnIleArgArgGluIleSerArgLysSerLysTrpValAspMe 81
Db 2576 -----ACAGCAATTTCGGCGGAGATCAGCCGAAGAAGCAAGTGGGTGATAT 2623
QY 81 tLeuGlyAspTrpGluLysTrpLysSerSerArgLys-----
Db 2624 GCTGGAGACTGGAGAAATACAAAAGACAGAAAAGTAACGTGTGAGGAGAGAGCA 2683
QY 93 -----
Db 2684 CTCTTCAGAGACAGGGGACAGGACCCATGCTTGTGCTGCGACCATCAGCTCTC 2743
QY 93 -----
Db 2744 AGAGGTGGGGCACACTGTCTCGCCAGAGACTGACGGCCTGTGCGCAGATTTC 2803
QY 93 -----
Db 2804 TGCTATTGTCGAAGCTGACCTTGACAGGAGGAAATCTGAATTAAGGCTGGACTAC 2863
QY 93 -----
Db 2864 CCGAGCTCAAGGCTAGGATGCTGTGTGACCTGAAGAGAAAGTTCAATCAGA 2923
QY 93 -----

Db 2924 GTTTCAGTCTGAGTGCATCCACTCTTTCAGTCTGGGAAGGAGACCCTGTCCAGC 2983
 QY 93 ----- 93
 Db 2984 TTGATCTCACCCTTACTGAGGAATCATGGGGCCAAAACCAATTTCCAGAAATCCCGG 3043
 QY 93 ----- 93
 Db 3044 GCTCTGGTCCCTCACTGGGGGTCACCCCGTGGCCCTGTGACACCAAGATTGTTTCTGCCACA 3103
 QY 94 -LeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerVa 113
 Db 3104 GCTCATAGATCGAGCATACAAAGGAATGCCATGAACTCCGGGCCCATGTGTGTCAGT 3163
 QY 113 IleuIleAsnIleGluGluMetLysLeuLysAsnProGlyArgTyrGlnIle----- 130
 Db 3164 CCTCCTGAACATTGAGGAATGAAGTTGAAAAACCCCGGAAGATACAGGT-ACGCTCAG 3222
 QY 130 ----- 130
 Db 3223 CCAGAGCACAAACAGAGCAGGCCGTGCGGGGCCAGGCTCCAGCTGGAAGGAAAG 3282
 QY 130 ----- 130
 Db 3283 TCAAGACCACCCTGGGAGCTGGGGGTGAAGGTCAATGAACACCCTGGGCAAGATGTT 3342
 QY 130 ----- 130
 Db 3343 GACACAGTCACACACAGAACAACTCAGCTGTGTGACCTCCCTGCTTCAATGAACAGCC 3402
 QY 130 ----- 130
 Db 3403 AAAATGCAAGCTTCTGAGAGAAGAACTTCCTCTGTCTCTCCGGAAGTGTGAC 3462
 QY 130 ----- 130
 Db 3463 TGTGGGCTGACTGCCACTGGGGGCAAGGAGTCTTCCATCTGTTCTGAGACTGCTTCTCC 3522
 QY 131 -----MetLysGlyLysGlyLysArgSerSerGluHisIleGlu 143
 Db 3523 GCTTGGCCCTGCCCTACAGATCATGAAGAGAGGGCAAGAGTCACTGAGCACATCCA 3582
 QY 143 nArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTy 163
 Db 3583 GCGCATCGACCGGACATTAAGCGGACATTAAGAACATATGTTCTTCAGAGGATCGATA 3642
 QY 163 rGlyThrLys----- 166
 Db 3643 CGGAACCAAGTAAGCCTACGGGAGCCACAGGGTCCACAGAGATGGGGTGAATGAGAG 3702
 QY 166 ----- 166
 Db 3703 GATGGGGGCTTCCCGGAGAGAGCAAGGCTACCCAGAGGAGTACACAGCTGCCAA 3762
 QY 166 ----- 166
 Db 3763 GAGCTCTCCCGGCCAGGAGAGCAGCCGACCATGAACCGAGCACCCTGCTTCCAAG 3822
 QY 166 ----- 166
 Db 3823 CCCTGGGCGAGACTGGAATGTGGGGCCAGAACCCAGAGGATCCTGAGAGATGGAAG 3882
 QY 166 ----- 166
 Db 3883 GCAGCAAAACAAATCATGCACAATGTGAAGGTGCTCTCCCTGACCCATGGGGACCCAT 3942
 QY 166 ----- 166
 Db 3943 GGTAAGACCCACGGGAGGTGGCAGGATAGAGGGCCCATGAGCCCCCAGGCAACAGTG 4002
 QY 166 ----- 166

Db 4003 ACAGCACCAAAATGCTGGGAGAAATTAGGGGTCTGAAACTCTCATCCAGGTCCGCTGGGA 4062
 QY 166 ----- 166
 Db 4063 ACATGACATGGCACAGCCACGTTGGCAGCAGTTGGCAGTGGCTGCAAAAGCTGATGG 4122
 QY 166 ----- 166
 Db 4123 ACTGAACACACATCCCAAAGTGTACAGATATTGAACCCACTGATTGGAAACTGAC 4182
 QY 166 ----- 166
 Db 4183 ATCCACATGAACACGACATGCCAGTTCACTGCTTGAATCTCCTGCTCAACACAGGAG 4242
 QY 166 ----- 166
 Db 4243 CTTTGGGAGCGGCTTCAACACGGGAATGGGGAGACAAAGGCTGTCTCCCTCAAAC 4302
 QY 166 ----- 166
 Db 4303 GGAAGACCAGTGAGAAAAGGAAACGAGCCGTGATGCCCGACGAACGTGGTGATCC 4362
 QY 166 ----- 166
 Db 4363 TAGATGCAATTTGTGAGGAGACAGAACCCCAATAAGCTACCAAGTAGATTCCC 4422
 QY 166 ----- 166
 Db 4423 ATTCCTAGGCCATTCTGAAAAAGGCCAAACACAGGACTGAGAGCAGTCTGGGTGGCC 4482
 QY 166 ----- 166
 Db 4483 AGGGGCTGACCGATCGGGGAGAGGCTGGGTGATAGGGGCCACCTGAGACTTGAGGA 4542
 QY 166 ----- 166
 Db 4543 TGAAGAGTCGCCCAAGAGGGGCTGAGCGGTGCGCGGAGACTCTGCACATTGTTTG 4602
 QY 166 ----- 166
 Db 4603 GAACCGTGAGAACTGTACACACACAGACTGAATGCGGTGTGTGCAAACTGAAAAAA 4662
 QY 166 ----- 166
 Db 4663 AAAAAATCATTCAGAGTGAAGAGATGAGCAAGTCACTGTACACTGGCTATTGTC 4722
 QY 166 ----- 166
 Db 4723 ATGTCAAGATGTGATTTTACTGAACAATTTCTCAAGAGTCTCAGGCCCTGAAGAGCT 4782
 QY 166 ----- 166
 Db 4783 CACTGCTTATCTGTGAACAATCTGAACCTGAATGGATTGCTTGAAGCTTTGTAGA 4842
 QY 166 ----- 166
 Db 4843 CAAAGTGAATTAACAACATCTGCACAAAACAAAGCCCTTCTCTGTTCTTA 4902
 QY 167 --GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLysAsnPro----- 182
 Db 4903 GGCAGCGGGAACCTACATCTCCTCGGCATATGAGAGTAAACCCGGTGAATATTC 4962
 QY 182 ----- 182
 Db 4963 CCGGCAGTGAGTTCGGGGCCATATTTCCATATTGACAGAGTGGGTCTGTGGGGG 5022
 QY 182 ----- 182
 Db 5023 TGTCTGTCTTTTAAAGTTAGTATTGTGACCCACAGGATATAGAGGTAGATGT 5082
 QY 182 ----- 182
 Db 5083 CAGCTACCGCTGGCATAAACCTCCAAAGAGGGGTGTCTCAAGGGGTCAAGCTGAGA 5142

QY 182 ----- 182
Db 5143 CACAAGAGATCAGGGCCCGACTCTGTGTCACTGGGCTGACCACTTCTCAGA 5202
QY 182 ----- 182
Db 5203 ACAAGAAATGACGCCCTCTCTGGGCTGCCAAAGCCCAAGAGCTTGGCAGCATCGC 5262
QY 182 ----- 182
Db 5263 ACACAGATGTGTCTATCAGCAGACATTTGACAAGGTGTGAAGTGCCTGATGACTT 5322
QY 182 ----- 182
Db 5323 GGCTCTTGTCATGAATAATGTGCACTGAGAAAGCTCTTTTTCAGAGAAAGCTCT 5382
QY 182 ----- 182
Db 5383 CCTTCAGAGAAAGCCTCTCCAGTCACTCTGCCCTCCAATGACATGAGTCTCCAGG 5442
QY 183 ----- 185
Db 5443 TGACCTCAGCCCTCCAGGTATGTCTTCCATGTGACTCTGGCTCTTGCAAGAGGTGG 5502
QY 185 |YTYrCyAaTgAspLeuSerHisIleAlaIleuPheLeuLeuTyrLeuProGluGlu 205
Db 5503 GCTACTGACGGGACCTGAGCCACATGCGCCCTTGTCTCTCTATCTTCTTGAGAGG 5562
QY 205 sPAlaPheTrpAlaLeuValGlnLeuLeuIleSerGluuArgHisSerLeuGln----- 222
Db 5563 ATGCATTCTGGGCACTGGTGCACTGTGGCAGTGAGAGGCACTCCCTGCA-ggGTAAG 5621
QY 222 ----- 222
Db 5622 TGAACAGCTGCCCCGGGACCTCTGCAAGCAGACCTGGGGATGGCCACCTGGCCGGGT 5681
QY 222 ----- 222
Db 5682 GATCACAGCTTTCAAGCCAAGGACCTCTTGTGTGCGCAGCTTGTGGAGACTTTAG 5741
QY 222 ----- 222
Db 5742 ATGTCTGTGTAGGGTCCCAAGAGTCCACGGCTGACCCCCCAAGCCCAATCAGAGC 5801
QY 222 ----- 222
Db 5802 CCTCTCATCCCATCAGCAGAGGGCATCTCATCTCCCGTGCCACCTCTGTGTCTG 5861
QY 222 ----- 222
Db 5862 GAGCCAGGCCCTCCGGCTGTATTCTGTGAGCTGACTCTCCCTCCCTGAGAGTCTCC 5921
QY 222 ----- 222
Db 5922 TGCCCTCCAGCTGCCCCGGGCTCTGTGCCATCGGTGCCAAGAATGGGCCCAAGCC 5981
QY 222 ----- 222
Db 5982 CAGGTGGCAGCATCTCCCATCTCCCTGTCTCCCTGCGCCGACCCCACTACCAAGAGATGA 6041
QY 222 ----- 222
Db 6042 CCGGGAAGCCCAAGCCCAAGCTTCCGGCGGCTGTGTGCTGAAGTCAAGGCTT 6101
QY 222 ----- 222
Db 6102 GCCCTTTTTCACCTTGCCCAAGAGGCTTCAGGGGAACCTCCAGCAGGCTCCAGGA 6161
QY 222 ----- 222
Db 6162 ATGTTCCCGCCCCACCTCCCAAGGTAAAGGCGCATGTTGGGTTCACCAAGTGGAGGG 6221

QY 222 ----- 222
Db 6222 TGGAGTAGCTTGGGGTTTGGGGCCCTCTCCAGCTGCCCAGCTTTCAGCTGATGCT 6281
QY 223 ----- 226
Db 6282 CCACATCTTGGGGGAAGGCTCTGATTTTCATGATGGCTGGGGCTTCTCAGGATTTCA 6341
QY 226 erProAengIyGlyThrValGlnGlyLeuGlnAspGlnGlnHisValAlaThrS 246
Db 6342 GCCCAATGGCGGGACCTGCCAGGGGCTCCAAAGACCAAGAGCATGTGTAGCCACGT 6401
QY 246 erGlnProLysThrMetGly----- 252
Db 6402 CACAATCCAAGACCATGGGGCATCAGGTGAGTTATGTCTCCCTCAGCTTCCCAAGAG 6461
QY 252 ----- 252
Db 6462 CCTGCCCTCCCGTGGGGCTGTAGAGCAGGGGGCTGGGGCCCTCGTGGGGCTGTGAC 6521
QY 252 ----- 252
Db 6522 TGGCTGAGTCCCAAGCAGGGGCTGACCTGGACGTGCGGTTCTCATTGGCTGGAGTTG 6581
QY 252 ----- 252
Db 6582 GTTTCCTTCTTGCCCTGAGGAGACAGAGGCAAGGATGGGGCCAGCTCCCGCAGA 6641
QY 252 ----- 252
Db 6642 GCAGGGCAAGGGCAGTGTGTCCACCGGAGTGTGGAAAGTGAAGTGTGTGGGGAGC 6701
QY 252 ----- 252
Db 6702 TCTGACACCGCCCAAGTCTTCTGACTAAGGGAAGGTCCTCAGAGACCCGAGAAGGG 6761
QY 253 ----- 258
Db 6762 AGGTTTTTAGGGCAGCCCAAGTGCCCTGAGCACCCTGTGCTTCCATCAGACAAGAA 6821
QY 258 sPLeuCyAsGlyGlnCySerProLeuGlyCysLeuIleArgIleLeuIleAsp-Gly--- 276
Db 6822 ATCTATGTGGGAGTGTTCCTCCGTTAAGCTGCTCATCCGATATGATGACGGGGTAA 6881
QY 276 ----- 276
Db 6882 GGAGGCATAGGGAAGACCTGTGCTCAGGACCTTCTTGCCCTGCAGTGCCCTGTCTCC 6941
QY 276 ----- 276
Db 6942 AGCCCGGGGTCTGGCTCACTCCAGCCCAAGAGGCTCAGGGGGTCCCAAGAGCA 7001
QY 276 ----- 276
Db 7002 CACAAGCAAAACCTTGTGCCCAAGAGGGTCAATCCAGGGCAATGGCTGGGCTCAGGCC 7061
QY 276 ----- 276
Db 7062 CAGCTCATGGGCACTGGGCCAGAACCGACTTGAGAGGGCTCAGGGAAGCTTAAGC 7121
QY 276 ----- 276
Db 7122 CTTGGGCAAGCCCTCTCTCCAGAGCCACATCCCACTCAATGATGCCCCCATGAG 7181
QY 276 ----- 276
Db 7182 GAGCTTCAAGACCTTGTCTGACCCAGCGCTCTGAGAGGGCTCAGGCGACCTCATGGGAA 7241
QY 276 ----- 276
Db 7242 GGTCACTGACTTGAAGACTGAAGCCCAAGTGTGCGAGCTCGAGCCACAGCCCCAGCC 7301
QY 277 ----- 283
Db -----IleSerLeuGlyLeuThrLeu 283


```
Db      7302  TGGAGGACCAGGTTCTTTCACACCTGCTGTCCCAACAGATCTCTTCGGGCTCACCCCTG  |||||
QY      284   ArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIleThrArgIle  |||||
Db      7362  CGCTGTGGGACGTGTATCTGTAGAAAGCGAACAGCGCTGATGCCATTAACAAGAATC  |||||
QY      304   AlaPheValGlnGln-----  |||||
Db      7422  GCCTTTAAGTTCAACAGATAAGTCTAGCTGTGCCAGCGGGCTGGGGAGCCCTGG  |||||
QY      309  -----  |||||
Db      7482  GTCAAGACCCGACTGCCCCGAGGCAAGCTCTCACACTGTCTCATGATCCGCTGTTCT  |||||
QY      309  -----
Db      7542  GGGCCAGAGGAGTCCGGCCAGGTGGGCTGGGCAAGACACTGTGAACCGAGCCCATCC  |||||
QY      309  -----
Db      7602  CTCACATGATCCGGACAGGGAAGTGTCAACCACTCTGACTTTGATCTGGGTCCCAAG  |||||
QY      309  -----
Db      7662  CCACAGTCTCCTGTGTATATCTGACACACTGGGGGTGGCCACAAAGATCCGGCACCGAC  |||||
QY      309  -----
Db      7722  CAGTAGAGACTGAAGTGGCCACGGGTAAGAGCTGTGACCATTCACAGTAACCTCCCT  |||||
QY      310  -----
Db      7782  GGCGTGATATCCACCTGTCCCTAGAGCGCTCACGAAGACGTCCAGGTGTGGCCCGTGG  |||||
QY      322   AlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLys  |||||
Db      7842  GCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGGACACTGTGTCTAAG  |||||
QY      342   HisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProAla  |||||
Db      7902  CATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACTGCCCCCAGG-  |||||
QY      361  -----
Db      7961  TGGGCTCCAGTGCATGTCCCTCCCATGTCAACCTCTGGGTAATCAAGTAGGGGAG  |||||
QY      361  -----
Db      8021  TGCCCGGGAACCGCAACCTACTACCTGGGCTTCTCTTCACTTTTCTCTCTCTT  |||||
QY      361  -----
Db      8081  CCTCTGACTTAAGAAAGTACAGAGGCCCCACCGGTCTCTCAGGGCAGGGCTCAGTGC  |||||
QY      361  -----
Db      8141  GTGTATACTGACATGCTGTGCAAGCAGAGGGGATGTGGGCAAGACCCTCAACAAGC  |||||
QY      361  -----
Db      8201  CCCCTCCCACTTTCACGGTGTCTCCCTGTCCCTCGCAGGGCCCTCCAAGTTACTAGA  |||||
QY      361  -----
Db      8261  CGAGCCCAAGACCAATTGTGGAGAACCCCGCCCTCCCTGCAAGCAACCAAGCCTCAGA  |||||
QY      361  -----
Db      8321  GAGCAGCAGAGGCCCTCACTCTGCAAGTCTCTCAAGGTTGCCAGGACAAGAGCCTG  |||||
QY      361  -----
```

```
Db      8381  GAGCCAGGAGACAAAGGAATCCGTGTCCCTGACCCACAGAGCATTCAGGAGAGGCC  |||||
QY      361  -----
Db      8441  AGAGCCAGAGCCAGAAGTTACGCCAAGAGTGGAAACGGTCACTCTGGCATGACTGGGC  |||||
QY      361  -----
Db      8501  AGCCAGAGGAGGAGAGGTGACCCACGTCCGGGCCCAATCAACCACTGCGAGACGGGT  |||||
QY      361  -----
Db      8561  CCCCACGTAGGTGACAAAGGGGCTGGGTGACATCCAAGGCCCTCCACCTGAGTTTGA  |||||
QY      361  -----
Db      8621  CTGGGGCCGTATCCAGGCCCAACAGCCCTGGGACGAAGTGTGGCAGGAAGCCCC  |||||
QY      361  -----
Db      8681  AGCCAGTGAACCTTGGGGCAGTCCCAAGAGCAACCCGCGATGCCACGACAGCTTCC  |||||
QY      361  -----
Db      8741  CACGCCAGGACATGACCCCTCCCTCTGGGATCAGCAGACTACAGCGTGTCTCGGT  |||||
QY      361  -----
Db      8801  GTCAAGGCCAGGGGGCCACACAGGGAACCCGAGACTCCGAGATGCAAGGAGTGGGGCC  |||||
QY      361  -----
Db      8861  CAGCCCGGAAGGCTTGCCTGGGCTCACTGAGATGCTGACCGCGTCTGTTCTTTCA  |||||
QY      362  ---LysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgLys  |||||
Db      8921  GCCAAACCCGAGCAAGGGTGTGTGGATTCAGGCTGTGCGGCTTCAAGTGGCGGAG  |||||
QY      381   ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro  |||||
Db      8981  ACCCTGTCAAGGGGAGACAGGAGGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC  |||||
QY      401   IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla  |||||
Db      9041  ATTTGGTCAAGTTCCTCCCGCAAGGACCTGTTCTTCCACACCTGTCTGTGGGCT  |||||
QY      421   ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln  |||||
Db      9101  GTCCGGGAAGACCTAACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG  |||||
QY      441   GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr  |||||
Db      9161  GGAGACCTCAGGGTCTCTGAGATTCTGAGTGAATCCATGCCCCGCTCCCAAG  |||||
QY      461   AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal  |||||
Db      9221  GACCTGACGTAGAGGGGCCCTGTGTTCCGCATATGATTCAGACAGAGCTGTGGTTC  |||||
QY      481   ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu  |||||
Db      9281  CGTGCCATATCCAGAGGAGACCAAGTGGCCCCCTGTGGCAGGCTGAACACCTGGGAG  |||||
QY      501   ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg  |||||
Db      9341  CGGGTGAATCGGCTTTCGTGCTGCAACCAAGCACTGATTCGACCAAGGCAACCCCTTCA  |||||
QY      521   AlaArgAspGlnGlnProCysAlaAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu  |||||
Db      9401  GCTAGGAGCAACAGCAGTGTGCTCCCACTCAGGGCCTTGCTGCGGCTTCACTTG  |||||
QY      541   GluSerSerGlnPheProProGlyPhe  |||||
Db      9461  GAAAGTCTCAGTTCCCTCCAGGCTTC  |||||
```

RESULT 14
US-09-764-891-7857
; Sequence 7857, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7857
; LENGTH: 9805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7857

Alignment Scores:
Pred. No.: 1.58e-128 Length: 9805
Score: 1507.50 Matches: 487
Percent Similarity: 20.97% Conservative: 6
Best Local Similarity: 20.71% Mismatches: 17
Query Match: 50.54% Indels: 1844
DB: 10 Gaps: 11

US-10-071-838-2 (1-549) x US-09-764-891-7857 (1-9805)

QY 24 LysGlyHisArgAlaGlyLeuProGluAspLysGly-----ProLysProPheArgSer 41
Db 2492 CAGGGCTCCAGAGCCCAAGACCAGCATCCCGGGGCTCTGGAACTCTGGCAGCTCC 111
QY 42 TyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeuThr 2551
Db 2552 GCTAACTCCACATGCTCTCATTTG-----
QY 62 AlaArgGluAlaLys-GlnIleArgArgGluIleSerArgLysSerLysTrpValAspMe 2575
Db 2576 -----ACAGCAAAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGGTGATAT 111
QY 81 tLeuGlyAspTrpGluLysTyrLysSerArgLys----- 2623
Db 2624 GCTGGAGACTGGAGAAATACAAAGACAGAAAGTAACTGTGGAGGAGAAACA 93
QY 93 ----- 2683
Db 2684 CTCTCTGACAGACAGGAGACAGCAACCATGGCTGTGGCCATCATGCTCTC 93
QY 93 ----- 2743
Db 2744 AGAGGTGGCGGCACACTGTCTCGCCAGAGACTGAGCCCTGTGCGCAGATTTC 93
QY 93 ----- 2803
Db 2804 TGCCTATTGTCGAAGCTCACCTTGAGAGGAGAAATCTGAATCTAGGGCTGGACTAC 93
QY 93 ----- 2863
Db 2864 CCGAGACTCAAGGCTAGGATGCCCTGTGACTGAAGAGAGAAAGTTGAGATCAGA 93
QY 93 ----- 2923
Db 2924 GTTTCGACTCTGAGTGTCCATCCACTCTTTCAGTCTGGAGAGGAGACCTGTCCAGC 93
QY 93 ----- 2983
Db 2984 TTGATCTCACTTACTGAGGATCATGGGGCCAAACCAACAATTTCAGATCCCGG 3043
QY 93 ----- 3043
Db 3044 GCTCTGTCTCACTGAGGTCAACCCCGTGGCTGTGACACAGATTTGTTCTGCCACA 3103

QY 94 -LeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTyrSerVa 113
Db 3104 GCTCATATGATGAGCATACAAAGGAAATGCCCATGAACATCCGGGGCCGATGTGTCACT 113
QY 113 tLeuLeuAsnIleGluGluMetLysLeuLysAsnProGlyArgTyrGlnIle----- 130
Db 3164 CTTCTGAACATTTAGGAAATGAAGTTGAAGAAACCCCGAGATACCAAGT-ACGCTCAG 3222
QY 130 ----- 3222
Db 3223 CCAAGACAAACAAACAGACAGGCCGTGTGGGGCCAGGTCTCCAGCTGAGGAAAG 130
QY 130 ----- 3282
Db 3283 TCAGAACCAACCTGGGGAGCTGGGGGTGAAGTCAATGAACACCTGGGACAGATGCT 130
QY 130 ----- 3342
Db 3343 GACACAGTACACACAGAAACTCAGCTCTGTGACCTCTGCTTCAATACAGACC 130
QY 130 ----- 3402
Db 3403 AAAATGACCTTTCTGACAGAAACCTTCTTCTGTCTTCTTCCGAAGTGTGAC 130
QY 130 ----- 3462
Db 3463 TGTGGCTGACTGCCACTGGGGGAGGAGTCTTCCATCTGTCTGAGACTGCTCTCC 130
QY 131 ----- 3522
Db 3523 GCTTGGCCCTGCTTACAGATCATGAAGAGAGGCAAGAGGTCTCTGACCATCTCC 143
QY 143 nArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTy 3582
Db 3583 GCGCATGACCGGACATTAAGCGGACATTAAGAGACATATGTTCTCAAGGATCGATA 163
QY 163 rGlyThrLys----- 3642
Db 3643 CGGAACCAAGTAAGCTTAAGGAGCCACAGGCTCCACAGATGGGGTGAATGAGAG 166
QY 166 ----- 3702
Db 3703 GATGGGGCTTCCCGAGACAGAAACAGGGTCAACCAGAGGATGACAGCTGCCAA 166
QY 166 ----- 3762
Db 3763 GAGCTTCCCGGCCAGGAGACAGCCGACCATGAACGAGACCTCTGTTCCAAG 166
QY 166 ----- 3822
Db 3823 CCCTGGCCAGACTGGAACATGTGGGGCCAGAACCCAGAGATCTTGAGAGATGGAAG 166
QY 166 ----- 3882
Db 3883 GCAGCAACAAATATGACAAATGTGAAGGTGCTTCTCTGACCCATGGGGACCAT 166
QY 166 ----- 3942
Db 3943 GGTAGACCAAGGAGGTGACAGATAGAGGCCCATGAGCCCCCAGCAACAGTG 166
QY 166 ----- 4002
Db 4003 ACAGACCAAAATGCTGGAGAAATTAAGGGTCTGAAACTTTCATCCAGTCCGCTGGA 166
QY 166 ----- 4062
Db 4063 ACATGACATGACACAGCAGTTGGACGAGTTGGCAGTGGCTCAACAAAGCTGATGG 166
QY 166 ----- 4122
Db 4123 ACTTGAACACACATCCCAAGTGTACAGATATTGAACCCACTGATTTGAAACTGAC 166

QY 166 ----- 166
Db 4183 ATCCACATGAAMCAGCATGCCAGTTCACTGCTGTACTCTCGTCACTCACACGAG 4242
QY 166 ----- 166
Db 4243 CCTTCGGGACGCGCTTCAACACGGGAATGGGGAGCAAGCTGCTCCCTTCAAC 4302
QY 166 ----- 166
Db 4303 GGAAGACCCAGTGAGAAAAGGAAACGAGCCGGTATGCCCGCACGAAGTGGGTGATCC 4362
QY 166 ----- 166
Db 4363 TAGATGCAATTTGCTGAGGGAAGAACCCAGACCAATTAAGCTACACAGTAGGATCCC 4422
QY 166 ----- 166
Db 4423 ATTCCTAGGCCATTTGGA AAAAGCCAAACCAAGGAGCTGAGAACAGTCTGGGTGCC 4482
QY 166 ----- 166
Db 4483 AGGGGCTGACGATCGGGGAGAGGCTGGGTGATAGGGGCCAACCTGGAGACTTGAGAGA 4542
QY 166 ----- 166
Db 4543 TGAAGAGTCGCCCCAGAGAGGGGCTGAGCGGTGCCGGAGACTTGACATTTGTTTG 4602
QY 166 ----- 166
Db 4603 GAACCGTGAGGAAGCTGTACACACAGACTGAAGTGGCGTGTGCAAACTGAAAAAA 4662
QY 166 ----- 166
Db 4663 AAAAAAATCATTCAAGTGAAGAAAGATCAGGCAAGTCACTGTACAAGTGGGCTATTGTC 4722
QY 166 ----- 166
Db 4723 ATGTCAAGATGTGATTTTACTGAAACATTTTCAAGAGTCTCAGGCCCTGAAGAGCT 4782
QY 166 ----- 166
Db 4783 CACTGCTTATCTGTGAACATCTGAACCTGAAATGGGATTTGCTGTAGGCTTTGTAGA 4842
QY 166 ----- 166
Db 4843 CAAAGTGAATTAACAACATCTGCACAAAACAAACCAAGCCCCCTTCTCTGTTCCCTA 4902
QY 167 --GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyrAsnPro----- 182
Db 4903 GGCAGCGGGAACCTACATCTCCTCGGTATGAGAGATATTAACCCGGTGAGTATTC 4962
QY 182 ----- 182
Db 4963 CCGGCAGTGAGTCCCGGGCCATATTTCCATAATTGACAGAGTGGGTCTGTGGGG 5022
QY 182 ----- 182
Db 5023 TGTCTGTCTTCTTTAAAGTATGATTTGTGACCCACAGATATAGAGGTAGATGT 5082
QY 182 ----- 182
Db 5083 CAGCTCACCGCTGGCATAAACCTCCAAGGAAGGGGGTGTCTCAAGGGGTCAAGCTGAGA 5142
QY 182 ----- 182
Db 5143 CACAAGAGAGTCAGGGCCCGGAAGCTCTGTGTGCTGAGCTGGGCTGACCACTTCTCAGA 5202
QY 182 ----- 182
Db 5203 ACAAGAAATGACGCCCTCTCTCTGGGGCTGCCCAAAGCCAGAGCTTGGCAGCATCGC 5262
QY 182 ----- 182

Db 5263 ACACAGATGCTGCTATCAGCAGACATTTTGAACAAGTGCTGAAGTGCCTGATGACTT 5322
QY 182 ----- 182
Db 5323 GGCTCTTGTCATGAATAATGAATGTGCATCCTGAGGAAGCCTCTTTTTCAGAGGAAGCCTCT 5382
QY 182 ----- 182
Db 5383 CTTTACAGGAAGCCTCTCCAGTCACTCTGCCCTCTCCAATGACATGAGTCTCCAGG 5442
QY 183 -----GluValG 185
Db 5443 TGACCTCAGCCCTCCAGGTGATGTCCTTCCATGCTGACTCTGGCTTTCAGGAGGTGG 5502
QY 185 1YTyrcysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyrLeuProGluGlu 205
Db 5503 GCTACTGCAAGGACCTGAGCCACATCGCGCTTGTCTCTCTATCTTCTGAGGAGG 5562
QY 205 spAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGln----- 222
Db 5563 ATGCATTCTGGGCACTGTGTGACAGTGTGCGCAAGTGAAGGCACTCTCTGCA-GGGTAAG 5621
QY 222 ----- 222
Db 5622 TGAACAGCTGCCCCGGGGAAGCTCTCTGACGACAGACCTGGGATGGCCACCTGGCGGGT 5681
QY 222 ----- 222
Db 5682 GATCACAGCTTTCAGCCAAAGCACCCCTCTGTGTGCGCAGCTTGTGGAGACTTTAGG 5741
QY 222 ----- 222
Db 5742 ATGTCTCTGCTGAGGTTCCACAGAGTCCAGGCTGACCCCAAAGCCCAATCAGAGC 5801
QY 222 ----- 222
Db 5802 CCTCTCATCCCCATCAGCAGAGGGCATCTCCTCCCCGTGGCCACCTCTGTCTCTG 5861
QY 222 ----- 222
Db 5862 GAGCCAAGCCCTCCGGCTCTGATTTGTGTGAGCTGACTCTCCCTCTGAGAGTCTCTCC 5921
QY 222 ----- 222
Db 5922 TGCCCTCAGTGCCTGGGCTCTCTGTCATCGGTGCCAGAAATGGGCCGACCAAGCC 5981
QY 222 ----- 222
Db 5982 CAGGTGGCAGCATCTCCCATCTCCCTGTTCCTCTGGCCGACCCCACTACAGAGATGA 6041
QY 222 ----- 222
Db 6042 CCGGGAAGCCCAAGCCCAAGTTCGGGCCGCCCTGTCTGCTGAAAGTCAAGGCTT 6101
QY 222 ----- 222
Db 6102 GCCCTTTTGCACCTTGCCCCAGAGGCTTCCAGGGGAACCTCCAGCCAGGCTCCAGGGA 6161
QY 222 ----- 222
Db 6162 ATGTCCGCCCCACCTCCCAAGGTAAAGCCGATGTTGGGTGACCAAGATGGAGGG 6221
QY 222 ----- 222
Db 6222 TGGGAGTAGCCTTGGGGTTTGGGGGCTTCCAGCTGCCAGCTTTCAGCTGATGCT 6281
QY 223 -----GlyPheHis 226
Db 6282 CCACATCTTGGGGGAAGCTCTGATTTGATGATGGGCTGGGGCTTCTCAGGATTTACAGA 6341
QY 226 erProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnValAlaIleThrs 246
|||||

Db 6342 GCCCAATGGCGGACCGTCCAGGGCTCCAAGACCAAGAGCATGTGTAGCCACGT 6401
QY 246 erGlnProLysThrMetGly----- 252
Db 6402 CACAAATCCAAAGACCATGGGGCATCAGGTGAGTTATGTGTCCCTCAGCTTCCCAAGAG 6461
QY 252 ----- 252
Db 6462 CCCTGCCCTCCCGTGGGGCTGTAGAGCAGGGGGCTGGGGCCCCCTCGTGGGCTGTGAC 6521
QY 252 ----- 252
Db 6522 TGGCTGAGTCCCAAGCCAGGCGCTGACCTGGGACGTGGGTTCTCCATGGGCTGGAGTTG 6581
QY 252 ----- 252
Db 6582 GTTTCCTTCTCTGCCCTGAGAGACAGAGGCACAGGGATGGGGCCAGCTCCGCAGA 6641
QY 252 ----- 252
Db 6642 GCAGGGCAAGGGCAGTGTGTCCACCGGAGTGTGGGAAGTGACAGTGTGTGGGAGC 6701
QY 252 ----- 252
Db 6702 TCTGACACCGCCAGTGTCTGCACTAGGGGAAGGTTCTTACAGACCCGAGAAGAGG 6761
QY 253 ----- 258
Db 6762 AGGTTTTTAGGCGAGCCCAAGTGGCCTGAGACCTCTGTGCTTCCATCAGACAGAAAG 6821
QY 258 sPLeuCySGlyGlnCySerProLeuGlyCySLeuIleArgIleLeuIleAsp-Gly--- 276
Db 6822 ATCTATGTGGGCAGTGTTCCTCCCGTGAAGCTCCTCATCCGATATTGATTGACGGGGTAA 6881
QY 276 ----- 276
Db 6882 GGAGGCATAGGAGACCTTGGCTCAGGGAACCTTCTTGCCCTGCAGTGCCCTGTCCCC 6941
QY 276 ----- 276
Db 6942 AGCCCGGGGCTGTGGCTACTCCCAAGCCCAAGAGGCTCAGGCGGTCCCAAGAGACA 7001
QY 276 ----- 276
Db 7002 CACAAGCAAAACCTCTGCCCCAAGGGGTCAATCCAGGCAATGGCTGGGGCTCAGGCC 7061
QY 276 ----- 276
Db 7062 CAGCCTCATGGGCGAGCTGGGCCAGGACCCGACTTGAGAGGGCTCAGGGAAAGCTCAAGC 7121
QY 276 ----- 276
Db 7122 CCTGGGCAAGCCCTCTCTTCCAGAGCCACATCCCACTCAAAATGAGTGCCCCCATGAG 7181
QY 276 ----- 276
Db 7182 GAGCTTCAAGACCTTGTCTGACCCAGCGTCTTGAGGGCTCAGGCGACCTCATGGGGAA 7241
QY 276 ----- 276
Db 7242 GGTCACTGACTGTGAGACTGAAGCCCAAGTGTGCGAGCTCGAGCCACCAAGCCCCAGCC 7301
QY 277 ----- 283
Db 7302 TGAAGGACCAAGGTTCTTTCACACCTGTCTGCCCAAGATCTCTCGGGCTCACCTTG 7361
QY 284 ArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIleThrArgIle 303
Db 7362 CGCCTGTGGGACGTGTATCTGTGAAGAGCGCAACAGCGCTTGATGCCGATAACAAGATC 7421
QY 304 AlaPheLysValGlnGln----- 309
Db 7422 GCCTTTAAGGTTCAAGACAGTAAGTCTACGTGTGCCACGCGGCTGGGGAACCTTGGG 7481

QY 309 ----- 309
Db 7482 GTCAGACCCCGACTGGCCGAGGGCAGCTTCTCTCACACTGTCTCATGATCCGCTTCT 7541
QY 309 ----- 309
Db 7542 GGCCGAGAGGAGTCCGGCCAGGTGGGCTGGGCGAGACACTGTGACACCGAGCCATCC 7601
QY 309 ----- 309
Db 7602 CTCACATGATCCCGACAGGAAGTCTCACACACTTTCGACTTTCATCTGGGTCCCCAG 7661
QY 309 ----- 309
Db 7662 CCAAGTCTCCTGTGTATATCTTGACACTGGGGTGCCACAAGATCCGGCACCGAC 7721
QY 309 ----- 309
Db 7722 CAGTAGAGACTGAAGTGGCCACGGGGTATGAGCTGTGACCATTTCCAGTTAACTCCCT 7781
QY 310 ----- 321
Db 7782 GGCCTGATATCCACCCTGTCTCTAGAGCGCCTCAGAAAGCTCCAGGTGTGCCCCGTG 7841
QY 322 AlaArgPheCySAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLys 341
Db 7842 GCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGCACTGTGCTCAAG 7901
QY 342 HisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProAla 361
Db 7902 CATCTTAGGGCTCTATGAAGAACTAAAGAAAGCAGGGGCACTGCCACCCCAAG- 7960
QY 361 ----- 361
Db 7961 TGGGCTCAGTGCCATGTCCCTCCCATGTCAACCTTGGGGTAGTCAGTAGTAGGGAG 8020
QY 361 ----- 361
Db 8021 TGCCCGGGAACCGCAACCTACTACCTGGGCTTCTCTTCACTTTCTTCTCTCTT 8080
QY 361 ----- 361
Db 8081 CCTCTGACTTAAAGAAAGTACAGAGGCCCAACGGTCTCAGGGCAGGCTCAGTGC 8140
QY 361 ----- 361
Db 8141 GTGTATCTGACATGCTGTGACGCGCAGAGGGGATGTGGCAAGACCTCCAACAAGC 8200
QY 361 ----- 361
Db 8201 CCCCTCCACTTTCCACGGTGTCTCTCTCTCCCTCGCAGGGCCCTCCAAGTTACTAGA 8260
QY 361 ----- 361
Db 8261 CGAGCCAGACCCATTGTGGGAGACCCCGCCCTCCCTGCAAGCACCCAAGCTCAGA 8320
QY 361 ----- 361
Db 8321 GAGCAGAGAGCCCTCACTCTGACGCTCTTCAAGTTGCCAGGACCAAGAACCTG 8380
QY 361 ----- 361
Db 8381 GAGCCAGGAGACAAGGAATCCGTGTCCCTGACCCACAGACATTCAAGGAGAGGGCCC 8440
QY 361 ----- 361
Db 8441 AGAGCCAGAGCCAAGATTCAAGCAGAAAGTGGGAACGGTCACTCTGGCATGACTGGGC 8500
QY 361 ----- 361
Db 8501 AGCCCAAGAGGGCAGAGGGTGAACCAAGTCCGGGCCCAATCAACCACTGCGGAGACGGGT 8560

QY 361 ----- 361
Db 8561 CCCACGTGAGTGACAAAGGGCTGTGTGACATCCAAGCCCCCTCCCACTGAGTTCTGA 8620
QY 361 ----- 361
Db 8621 CTGGGGCCGTATCCCAAGCCCCAACAGCCCTGGGACGAAGTGTGTGGAAGAAGCCCC 8680
QY 361 ----- 361
Db 8681 AGCCAGTCTGAACCTGTGGGGGCAAGTCCAGAGACCACCGCCATGCCACGACAGCTTCCC 8740
QY 361 ----- 361
Db 8741 CACGCCAGGACGATGACCCCTCTCTGGATCAGCAGACTACAGGCGTGTCTCGGT 8800
QY 361 ----- 361
Db 8801 GTCAGGCCACGGGGCCACACAGGAGCCCGAGACTCCGAGATGACGAGGAGGTGGGCC 8860
QY 361 ----- 361
Db 8861 CAGCCCCGAAAGCCTGCGTGGGCTCACTGAGATGCTGACCGCGTCTTTCTTCA 8920
QY 362 ---LysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 8921 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCGGCTTACAGTGGCGGGAAG 8980
QY 381 ThrLeuCySLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 8981 ACCCTCTGCAAGGGGACAGGACAGGCCCTCCAGGCCACCAGCCCGGTTCCCGCGCCC 9040
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySProGlyGlyAla 420
Db 9041 ATTTGTGAGCTTCCCGCCACAGGCGCACCTCGTCTTCCACACACCTGCTGTGGGCT 9100
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 9101 GTCCGGGAAGACACCTACCTGTGGGACATCAGGGTGTGCCAGCCGGCCCTGGCTCAG 9160
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 9161 GAGGACCTCAGGGTTCCTGAGATTCTCTGAGTGAATCCATGCCCTCCCAAGC 9220
QY 461 AspleuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 9221 GACCTGACGTAGAGGCGCTTGTTCCGCCATTATGATTGACAGACAGCTGTGGGTC 9280
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 9281 CGTGCCATATCCAGGAGGACCAAGCTGAGCCCTGCTGGCAGGCTGAACACCTGCGAG 9340
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 9341 CGGGTGAGATCGGCTTTCGCTGACACCAAGACATGATCCGACAGGACCCCTTCAAG 9400
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 9401 GCTAGGAGCAAGACAGCAGTGTCTCCCACTCAGGCGCTTGCTCTGCGGCTCCACTTG 9460
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 9461 GAAAGTTCTCAGTTCCTCCACAGGCTTC 9487

RESULT 15
US-10-091-504-1824
; Sequence 1824, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504

; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1824
; LENGTH: 9805
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-091-504-1824

Alignment Scores:
Pred. No.: 1.58e-128 Length: 9805
Score: 1507.50 Matches: 487
Percent Similarity: 20.97% Conservative: 6
Best Local Similarity: 20.71% Mismatches: 17
Query Match: 50.54% Indels: 1844
DB: 14 Gaps: 11

US-10-071-838-2 (1-549) x US-10-091-504-1824 (1-9805)
QY 24 LysGlyHisArgAlaGlyLeuProGluAspLysGly-----ProLysProPheArgSer 41
Db 2492 CAGGGCTCCAGAGGCCCAAGACCAGACATCCGCGGCGGCTGTGGAAAGCTTGACAGCTCC 2551
QY 42 TyrAsnAsnValAspHisLeuGlyLeuValHisGluThrGluLeuProProLeuThr 61
Db 2552 GCTAACTCCACATGCTCATTTG----- 2575
QY 62 AlaArgGluAlaLys-GlnIleArgArgGluIleSerArgLysSerLysTrpValAspMe 81
Db 2576 -----ACAGCAAAATTGCGCGGAGATCAGCCGAAAGCAAGTGGTGATAT 2623
QY 81 tLeuGlyAspTrpGluLysTyrLysSerSerArgLys----- 93
Db 2624 GCTGGAGACTGGAGAAATACAAAGACAGAAAGTAACTGTGAAGGAGGAACA 2683
QY 93 ----- 93
Db 2684 CTCTCTGACAGACAGGGGACAGGACCAATGCTTGTGCGCTGGCAACATCAGCCTCTC 2743
QY 93 ----- 93
Db 2744 AGAGGTGGGGGCACTGTCTCGCCCAAGAGACTGACAGCGCTGTGCCAGATTTC 2803
QY 93 ----- 93
Db 2804 TGCCTATTCTGCAAGCGTCACTTGCAGGGAGGAATCTGAATCTAGGGCTGGACTAC 2863
QY 93 ----- 93
Db 2864 CCGAGCTCAAGGCTAGGGATGCCCTGTGACCTGAAGGAAGAAAGTTCAAGATCAGA 2923
QY 93 ----- 93
Db 2924 GTTTCGACTGTGAGTGTCCATCCACTCTTCAATCCTGGGAAGGAGAACCTGTCCAGC 2983
QY 93 ----- 93
Db 2984 TTGATCTACCTCTACTGAGGATCATGGGCGCAAAACCAAAATTCCAGAATCCCGG 3043
QY 93 ----- 93
Db 3044 GCTCTGCTCTCACTGGGGTCAACCCCGTGGCTGTGACACCAAGATTGTTTCTGCCACA 3103
QY 94 -LeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyPrometTrpSerVa 113
Db 3104 GCTCATAGATCGAGCATACAAAGGAATGCCCATGAACATCCGGGCGCATGTGTAGT 3163
QY 113 IleuLeuAsnIleGlnGluMetLysLeuLysAsnProGlyArgGlyGlnIle----- 130
Db 3164 CTTCTGAACATGAGAAATGAAGTTGAAGAAACCCCGGAAGATACAGGT-ACGCTCAG 3222
QY 130 ----- 130

Db 3223 CCAAGCACACAAAACAGACAGGCCGTGCGGGCCCAAGTCTCCAGCTGAGGGAACG 3282
QY 130 ----- 130
Db 3283 TCAAGACCAACCTGGGGAGCTGGGGGTGAAGTCAGATGAACACCCTGGGACAGATGCT 3342
QY 130 ----- 130
Db 3343 GACACAGTCAACACAGCAAACTCAGCTGTGTGACCTCTCCCTGCTTCAATAACAAGCC 3402
QY 130 ----- 130
Db 3403 AAAATGCAAGCTTCTGCAAGAGAAACCTTCTTGTCTCTTCTTCCGAAAGTGTGAC 3462
QY 130 ----- 130
Db 3463 TGTGGGCTGACTGCCACTGGGGGCAAGGAGTCTTCCATCTGTTCTGAGACTGCTTCTCC 3522
QY 131 -----MetLysGluLysGluLysArgSerSerGluHisIleGlu 143
Db 3523 GCTTGGCCCTGCCCCCTACAGATCATGAAGAGAGAGGCAAGAGGTCACTGTAGACACATCCA 3582
QY 143 MArgLLeaspArgaspValSerGluYThrLeuArgLysHisIlePhePheArgAspArgTyr 163
Db 3583 GCGCATCGACCGGACATTAAGCGGACATTAAGGAAGCATATGTTCTTCAAGGATCGATTA 3642
QY 163 rGlyYThrLys----- 166
Db 3643 CGGAACCAAGTAAGCCTTACGGGAGCCACAGGGTCCACAGACAGATGGGGTGAATGAGAG 3702
QY 166 ----- 166
Db 3703 GATGGGGGCTTCCCGGAGACAGAAAGCCAGGGTCAACCAAGAGGATGACACAGCTGCCAA 3762
QY 166 ----- 166
Db 3763 GAGCTCTCCCGGCCACGGGAGCAGCGGCCACCATGAACCGAGCACCTCCCTGGTTCCAAG 3822
QY 166 ----- 166
Db 3823 CCCTGGGCCAGACTGGAACATGTGGGGCCAGAACCCAGAGAGATCCTGAAGAGATGAAG 3882
QY 166 ----- 166
Db 3883 GCAGCAAAACAATCATGCACAATGTGAAGGTGCTCTCCCTGACCCATGGGAGCCCAT 3942
QY 166 ----- 166
Db 3943 GGTAGACCCACGGGAGGTGGCAGATAGAGGGCCCATGAGCCCCCAGGCAACAGTGG 4002
QY 166 ----- 166
Db 4003 ACAGCACCAATGCTGGGAGAATTAGGGGTCTGGAAACTTCATCCAGGTCCGCTGGGA 4062
QY 166 ----- 166
Db 4063 ACATGACATGGCACAGCCAGCTTGGCAGCCAGTTGGGCAGTGGCTCACAAAGCTCGATGG 4122
QY 166 ----- 166
Db 4123 ACTTGAACCAACATCCCCAAAGTGTCAGATATTGAACCCACTGATTGGAAACTGAC 4182
QY 166 ----- 166
Db 4183 ATCCACATGAACCAAGCATGCAAGTTCACTGTTGACTCTCTGTCATCACAACGAG 4242
QY 166 ----- 166
Db 4243 CCTTCGGGAGCGCCTTCAACAAGGGAATGGGAGAGCAAGGCTGTCTCTCCCTTCAAAAC 4302
QY 166 ----- 166

Db 4303 GGAAGACCAAGTGAAGAAAAGGGAACGAGCCGGTGAATGCCCGCACGAACGTGGGTGATCC 4362
QY 166 ----- 166
Db 4363 TAGATGATTTTGTGAGGGAAGAGCCAGACCCCAATTAAGCTTACACAGTAGATTCCC 4422
QY 166 ----- 166
Db 4423 ATTCTAGGCCATTCTGGAAGGCAAGCCAAACCAAGGACTGAGAGAGCAGTGTGGTGGCC 4482
QY 166 ----- 166
Db 4483 AGGGCTGACGATCGGGGAGAGGCTGGGTGATAGGGGCCACCTGAGACTTGGAGGA 4542
QY 166 ----- 166
Db 4543 TGAAGAGTCGCCCCAGAGGGGCTGAGCGGTGGCCGGGAGACTCTGCACATTGGTTTG 4602
QY 166 ----- 166
Db 4603 GAACCGTGAGGAACGTGACACACAGACTGAATCGCGCTGTGTGCAAACTGAAAAAAA 4662
QY 166 ----- 166
Db 4663 AAAAAAATCATTCAGAGTGAAGAAAGATCAGGCAAGTCACTGTACAACTGGGCTATTGTC 4722
QY 166 ----- 166
Db 4723 ATGTACAGATGTGATTTTACTGAACATTTCTTCAAGAGTCTCAGGCCCTGAAGAGCT 4782
QY 166 ----- 166
Db 4783 CACTGCTTATCTGTGAACAATCTGAACCTGAATGGGATTGTGTAGCTTTGTAGA 4842
QY 166 ----- 166
Db 4843 CAAAGTGAATTAACAACATCTGCACAAACAAACCAAGCCCCCTTCTCTGTTTCCCTA 4902
QY 167 --GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluIuTyrAsnPro----- 182
Db 4903 GGCAGCGGGAACCTACTCCACATCTCTCTGTCATATGAGAGGTATAACCCGGTGAGTATTC 4962
QY 182 ----- 182
Db 4963 CCGCAGTAGAGTTCGCCGGCCATATTTCCATATTGACAGAGTGGTGTGTGGGG 5022
QY 182 ----- 182
Db 5023 TGTGTTGCTTCTTTTAAAGTTAGTATTGTGACCCACAGATATAGAGGTAGATGT 5082
QY 182 ----- 182
Db 5083 CAGCTCACCGGTGCATTAACCTCCAAAGAGGGGTGTCTCAAGGGGTCAAGCTGAGA 5142
QY 182 ----- 182
Db 5143 CACAAAGAGTCAAGGGCCCGGACTCTGTGTTCACCTGGGCTGAACCACTTCTCAGA 5202
QY 182 ----- 182
Db 5203 ACAAGAAATGACGCCCTCTCTCTGGGGCTGGCCCCAAAGCCAGAGACTTGCAGCATCCG 5262
QY 182 ----- 182
Db 5263 ACACAGATGTGTCTATCAGACACATTTTGGACAAGGTGTGAAGTGCCTGATGACTT 5322
QY 182 ----- 182
Db 5323 GGCCTTGTCTATGAATGAATGTGATCTGTAGGAAGCTCTTTTTCAGAGGAAGCTCT 5382
QY 182 ----- 182
Db 5383 CTTTCAAGGAAGCCTCTCCAGTCACTCTGCCCCCTTCCAATGACATGATGCTCTCCAGG 5442

QY 183 -----GluValG 185
Db 5443 TGACCTCAGCCCTCCAGGTGATGTCCTTCCATGGTGACTTGCGCTCTTGCAAGAGGTGG 5502
QY 185 lYTyCybArgAspLeuSerHisIleAlaIleuPheLeuTyrLeuProGluGluA 205
Db 5503 GCTACTGCAAGGAGCTGAGCCATCGCCGCTTGTCCTCTCTATCTTCTGAGAGAG 5562
QY 205 sPaIaPheTrPaIaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGln----- 222
Db 5563 ATGCAATTCTGGGCACTGTGCACTGCTGGCCAGTGAGGCACTCCCTGCA-GGGTAAG 5621
QY 222 ----- 222
Db 5622 TGAACAGCTGCCCCGGGACCTCTGACGCAAGACCTGGGGATGGCCACCCTGGCCGGGT 5681
QY 222 ----- 222
Db 5682 GATCAAGCTTTCAAGCAAGGCACTCTGTGTGCCAGCTTGTTGGGAGACTTTAGG 5741
QY 222 ----- 222
Db 5742 ATGTCTCTGTAGGGTCCCAAGAGTCCACGGCTGACCCCAAGCCCAATCAGAGC 5801
QY 222 ----- 222
Db 5802 CCTCTCATCCCATCAGCAGAGGGCATCTCATCTCCCGTGGCCACCCTGTGTCTG 5861
QY 222 ----- 222
Db 5862 GAGCCAGCCCTCCGGCTGTGATTCTGTGAGTGACTCTCCCTCCGTAGAGTCTCTCC 5921
QY 222 ----- 222
Db 5922 TGCCCTCAGCTGCCCGGCTCTGCTGCCATCGGTGCCCAAGATGGGCCCAAGACC 5981
QY 222 ----- 222
Db 5982 CAGTGGCAGCATCTCCCATCCCTGTCTCCCTGCGCCGACCCCATACAGAGATGA 6041
QY 222 ----- 222
Db 6042 CCGGGAAGCCAGCGCCCAAGTTCGCGCCGCTGTGCTGAAAGTCAAGGCTT 6101
QY 222 ----- 222
Db 6102 GCCCTTTTGCACCCCTGGCCAGAGGCTCCAAGGGAACCTCAGCCAGGCTCCAGGA 6161
QY 222 ----- 222
Db 6162 ATGTTCCCCGCCCCACCTCCCAAGGTAAAGCCGCATGTTGGGTCACCAATGGAGGG 6221
QY 222 ----- 222
Db 6222 TGGGAGTAGCCTTGGGGTTGGGGGCTCTCCAGCTGCCCAAGCTTTGCAGCTGATGGCT 6281
QY 223 -----GlyPheHis 226
Db 6282 CCACATCTTGGGGAGGCTGTGATTTCATGATGGGCTGGGGGCTTCTCAGGATTTTACA 6341
QY 226 exProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnIleValValAlaThrS 246
Db 6342 GCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGCATGTGTAGCCACGT 6401
QY 246 exGlnProLysThrMetGly----- 252
Db 6402 CACAAATCAAGACCATGGGCATCAGGTGATTATGTCCTCAGCTCTTCCAGAGG 6461
QY 252 ----- 252
Db 6462 CCTGCTCCTCCGCTGGGGCTGTAGAGCAGGGGGGCTGGGGCCCTCGTGGGGCTGTGAC 6521

QY 252 ----- 252
Db 6522 TGCGTAGTCCAGCAGGCGCTGACCTGGGAGTCGGGTTCTCCATGGGCTGGAGTTG 6581
QY 252 ----- 252
Db 6582 GTTTCCTTCTCGCCCTGAGAGAGACAGAGGCAAGGATGGGGCCAGCTCCCGCAGA 6641
QY 252 ----- 252
Db 6642 GCAGGCAAAAGGCGAGTGTCTCCACCGGAGTGTGGGAAGGTGACAGTGTGTGGGAGC 6701
QY 252 ----- 252
Db 6702 TCTGACACCGCCAGTGTCTGCACTAGGGGAAGGCTTTCAGAGACCCGAGAGAGGG 6761
QY 253 -----HisGlnAspLysBlyAs 258
Db 6762 AGGTTTTAGGCGAGCCCAAGTGGCCTGAGCACTCTGTGCTTCCATCAGACAGAAAGAG 6821
QY 258 sPLeuCybGlyGlnCybSerProLeuGlyCybLeuIleArgIleLeuIleAsp-Gly--- 276
Db 6822 ATCTATGTGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATGATTGACGGGTAA 6881
QY 276 ----- 276
Db 6882 GAGAGCATAGGAGACCTGGCTCAGGACCTTCTTGCCCTGCAAGTCCCTGCTTCCCC 6941
QY 276 ----- 276
Db 6942 AGCCCGGGGCTGTGGCTCACTCCAGCCCAAGAGGCTCAGGCGGGTCCCAAGAGACA 7001
QY 276 ----- 276
Db 7002 CACAAGCAAAACCTCTGCCCCAAGAGGGGTGATCCAGGGCAATGGCTGGGGCTCAGGCC 7061
QY 276 ----- 276
Db 7062 CAGCTCATGGCAGACTGGGCCAGACCCGACTTGAGAGGCTCAGGGAAAGCTCAAGC 7121
QY 276 ----- 276
Db 7122 CCTGGGCAAGCCCTCTCTCCAGAGCCACATCCCACTCAATGATGCCCCCATGAG 7181
QY 276 ----- 276
Db 7182 GAGCTTCAAGACCTTGTCTGACCCAGCGTCTGAGGGCTCAGGCGACCTCATGGGAA 7241
QY 276 ----- 276
Db 7242 GGTCACTGACTGTGAGACTGAAGCCCAAGTGTGCGCAGCTGAGCCACGACCCAGCC 7301
QY 277 -----IleSerLeuGlyLeuThrLeu 283
Db 7302 TGAAGGACCAAGTCTTTCACACCTGCTGTGCCACAGATCTCTCGGGCTCACCCCTG 7361
QY 284 ArgLeuTrpAspValTyrLeuValGlyGlyGlnAlaLeuMetProIleThrArgIle 303
Db 7362 CGCCTGTGGACGTGTATCTGTAGAAAGGGAACAGGCGTTGATGCGGATTAACAAGATC 7421
QY 304 AlaPheLysValGlnGln----- 309
Db 7422 GCCTTTAAGTTTCAGCAGAGTAAGTCTACGTGTGCCAAGCGGGGCTGGGAGCCCTGGG 7481
QY 309 ----- 309
Db 7482 GTTCAGACCCCGACTGGCCCGAGGGCAGCTTCTCACTGTCTCATGATCCGCTGTTCT 7541
QY 309 ----- 309
Db 7542 GGCCAGAGGAGGTCCGGCCAGGTGGGCTGGGCAAGACACTGTGACACCGAGCCCATCC 7601
QY 309 ----- 309

```
Db      7602 CTCACATGATCCGGACAGGGAAGTGTCTCACCACTCTCGACTTTCATCTGGGTCCCGAG 7661
QY      309 -----
Db      7662 CCACAGTCTCTGTGTATATCTGGACACCTGGGGTGGCCACAAGATCCGGACCGAC 7721
QY      309 -----
Db      7722 CAGTAGAGACTGAAGTGGCCACGGGTATGAGCTGTGACCAATCCAGTAATCCCT 7781
QY      310 -----
Db      7782 GGCCTGATATCCACCCTGTCTCTAGAGCGCTCACGAAAGCTCCAGGTGTGCGCCGTG 7841
QY      322 AlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLys 341
Db      7842 GCACGTTTTCGACACCGGTTCTGTGATACCTGGCCAGGGATGAGGACACTGTGCTCAAG 7901
QY      342 HisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProAla 361
Db      7902 CATCTTAGGGCTCTATGAAGAACTAAAGAAAGCAGGGGACCTGCCACCCCGAGG- 7960
QY      361 -----
Db      7961 TGGGCTCAGTGCATGTGCCCTCCCTCATGTCACTCTGGGTAGTCAAGTAGGGAG 8020
QY      361 -----
Db      8021 TGCCCGGACCCGCAACCTTAATACTGCGGCTTCTCTTCACTTTCTCTCTCTT 8080
QY      361 -----
Db      8081 CCTCTGGACTTAAGAAAGTACAGAGAGCCCAACGGTCTTCAAGGCGCTCAGTGC 8140
QY      361 -----
Db      8141 GTGTATACTGACATGTCTGTGCACCGAGAGGGGATGTGGCAAGACCTCCAAACAAGC 8200
QY      361 -----
Db      8201 CCCCTCCCACTTCCACGGTGTCTCTCTCTCCCTCGCAGGGCCCTCCAAGTTACTAGA 8260
QY      361 -----
Db      8261 CGAGCCCAAGCCATTGTGGGAGACCCCGCCCTCTCTGCAAGCACCAAGCCTCAGA 8320
QY      361 -----
Db      8321 GAGCAGCAGAGCCCCCTCACTCTCTGCACGCTCTTCCAAGTTGCGAGACAAGAGCCTG 8380
QY      361 -----
Db      8381 GAGCCAGGAGACAGAGGAATCCGTGTCTGTACCCACAGAGCATTCAGGGAGAGGGCCC 8440
QY      361 -----
Db      8441 AGAGCCAGAGCCAAAGATTCAAGCAAGTGGGAACGTCAGTCTTGCACTGGAGCTGGGC 8500
QY      361 -----
Db      8501 AGCCCAAGAGGCGCAGAGGGTGACCCAGCTCCGGGCCAAATCACCCACTGCCGAGACGGGT 8560
QY      361 -----
Db      8561 CCCCAAGTGAAGTGAACAAGGGCTGGGTGACATCCAAGGCCCTCCCACTGAGTTCTGA 8620
QY      361 -----
Db      8621 CTGGGGGCGGTATCCCAAGGCCCAAGCCCTGGGACGAAGTGTGTGCGAGGAAGCCCCC 8680
QY      361 -----
```

```
Db      8681 AGCCAGTCTGAACCTTGGGGGCACTCCAGAGGCCAACCCGCATGCCACAGACAGTTCCC 8740
QY      361 -----
Db      8741 CACGCCAGGACATGACCCCTCTCTTGGATCAGAGACTTACAGCGGTGTCTCGGT 8800
QY      361 -----
Db      8801 GTACGGCCACGGGGCCACACAGGAGCCCGAGACTCCGAGATGACAGGAGTGGGGCC 8860
QY      361 -----
Db      8861 CAGCCCGAAAGGCTGCGTGGGCTCACTGAGATGCTGACCGGCTGTCTTCTTTCA 8920
QY      362 ---LysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db      8921 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCACTGCGCGGAG 8980
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      8981 ACCCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGCCACAGCCCGGTCTCCCGCGGCC 9040
QY      401 IleTrpSerAlaSerProProAlaAlaProArgSerSerThrProCysProGlyGlyAla 420
Db      9041 ATTTGTACGCTTCCCGCCAGCGGCACTCTCTTCCACACCTGTCTCTGTGGGCT 9100
QY      421 ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db      9101 GTCCGGAGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 9160
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db      9161 GGAGAGCTCAGGTTCTCTGGAATTCTTCAGTGAATCCATGCCCCCGCTCCCAAG 9220
QY      461 AspLeuAspValGluGlyProTrpPheArgHisTyraAspPheArgGlnSerCysTrpVal 480
Db      9221 GACCTGAGACGTAGAGGCCCTTGTGTTCCGCAATTAATTAATTAATTAATTAATTAAT 9280
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db      9281 CGTGCATATCCAGAGAGACCAAGCTGGCCCCCTGTGCGCAGGCTGAACACCCCTCGAG 9340
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      9341 CGGTGAGATCGCTTTCGCTGCACCCAGCACTGAATTCGACCAAGGCAAGCCCTTCAGA 9400
QY      521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db      9401 GCTAGGAGACGAACAAGAGTGTCTCCACCTCAGGGCCTTGCTGCGGCTCCACTTG 9460
QY      541 GluSerSerGlnPheProProGlyPhe 549
Db      9461 GAAAGTTCTCAGTTCCTCCAGGCTTC 9487
```

Search completed: February 4, 2005, 13:50:57
Job time : 991 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 08:05:19 ; Search time 136 Seconds
(without alignments)
2869.287 Million cell updates/sec

Title: US-10-071-838-2
2983
Sequence: 1 MDVVEVAGSMWAGREDIIM.....TSGPCLGHLHSSQFPFPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10071838/runat_03022005_071219_215/app_query.fasta_1.711
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOBM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071838 @CGN_1_1_69 @runat_03022005_071219_215 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DEPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2850	95.5	2084	4	US-09-799-451-692 Sequence 692, App
2	2827	94.8	2072	4	US-09-799-451-691 Sequence 691, App
3	2133.5	71.5	8201	1	US-08-253-155A-9 Sequence 9, Appli
4	344	11.5	1981	4	US-09-620-312D-715 Sequence 715, App
5	337	11.3	1823	4	US-09-774-528-336 Sequence 336, App
6	258	8.6	3039	4	US-09-620-312D-675 Sequence 675, App
7	253	8.5	2955	4	US-09-620-312D-676 Sequence 676, App
8	238	8.0	1083	4	US-09-248-796A-5457 Sequence 5457, Ap
9	222	7.4	1887	4	US-09-270-767-12482 Sequence 12482, A
c 10	194.5	6.5	806	4	US-09-270-767-9899 Sequence 9899, Ap
11	182.5	6.1	4039	1	US-08-363-300-1 Sequence 1, Appli
12	172	5.8	1056	4	US-09-248-796A-1133 Sequence 1133, Ap

13	154	5.2	1659	4	US-09-248-796A-1820	Sequence 1820, Ap
c 14	149	5.0	7527	4	US-09-252-991A-71	Sequence 71, Appl
c 15	148.5	5.0	3763	4	US-09-919-039-243	Sequence 243, App
c 16	148.5	5.0	5185	4	US-09-976-594-640	Sequence 640, App
17	147.5	4.9	723	4	US-09-270-767-1491	Sequence 1491, Ap
18	147.5	4.9	723	4	US-09-270-767-16773	Sequence 16773, A
c 19	146.5	4.9	5037	3	US-09-428-517-1	Sequence 1, Appli
c 20	146	4.9	2304	4	US-09-799-451-340	Sequence 340, App
c 21	146	4.9	3102	4	US-09-252-991A-4429	Sequence 4429, Ap
c 22	146	4.9	43280	2	US-08-804-227C-1	Sequence 1, Appli
23	145.5	4.9	1368	4	US-09-252-991A-64	Sequence 64, Appl
24	145.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
25	145	4.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 26	144.5	4.8	2009	4	US-09-646-403-1	Sequence 1, Appli
c 27	144.5	4.8	12249	4	US-09-266-965-74	Sequence 74, Appl
c 28	144.5	4.8	18331	4	US-09-266-965-96	Sequence 96, Appl
c 29	143	4.8	1421	3	US-08-943-731-207	Sequence 207, App
c 30	143	4.8	2363	4	US-09-818-780-22	Sequence 22, Appl
c 31	143	4.8	6085	3	US-09-029-603-4	Sequence 4, Appli
c 32	143	4.8	20084	3	US-08-943-731-5	Sequence 5, Appli
c 33	142.5	4.8	990	4	US-09-266-965-18	Sequence 18, Appl
c 34	142.5	4.8	1593	4	US-09-252-991A-14376	Sequence 14376, A
c 35	142	4.8	1788	4	US-09-252-991A-4632	Sequence 4632, Ap
c 36	142	4.8	2250	4	US-09-252-991A-4181	Sequence 4181, Ap
37	142	4.8	5184	4	US-09-845-583A-9	Sequence 9, Appli
38	142	4.8	5184	4	US-09-561-709B-4	Sequence 4, Appli
39	142	4.8	8037	4	US-09-774-528-209	Sequence 209, App
c 40	141	4.7	34094	4	US-09-292-034-1	Sequence 1, Appli
c 41	140	4.7	827	4	US-09-620-405B-486	Sequence 486, App
c 42	140	4.7	827	4	US-09-604-287A-486	Sequence 486, App
c 43	140	4.7	827	4	US-09-834-759-486	Sequence 486, App
c 44	140	4.7	1518	4	US-09-620-405B-484	Sequence 484, App
c 45	140	4.7	1518	4	US-09-604-287A-484	Sequence 484, App

ALIGNMENTS

RESULTS

US-09-799-451-692
Sequence 692, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_pl_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

LOCATION: (163) ..(1755)
US-09-799-451-692

Alignment Scores:
Pred. No.: 2.24e-201
Score: 2850.00
Percent Similarity: 96.36%
Best Local Similarity: 96.17%
Query Match: 95.54%
DB: 4
Length: 2084
Matches: 528
Conservative: 1
Mismatch: 2
Indels: 18
Gaps: 1

US-10-071-838-2 (1-549) x US-09-799-451-692 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGCTGGTAGAGCTCCGGGCGAGTTGGGCGACAAAGACGAGAGGACATCATTATG 222
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysBProPheArg 40
Db 223 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGGACAGGGCCCTTAACCTTTTGA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheArg 60
Db 283 AGCTACAAACAACAGCTCGATCATTTGGGATTGTACATGAGACGGAGCTCCCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluLysSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGGAT 402
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGAGAAATACAAAGACAGCAAGATCATAGATCGAGCTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAAGAGCATATATTCTTCAGG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
Db 589 GATCGATACGGAACCAAGACCGGGAATCTCCACATCTCTGCGCATATAGAGAGTAT 648
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGGCTACTGACGGAGCTGAGCCACATCGCCGCTTCTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGGAGATGATCTTGGGCACTGTGACGCTCTGCGCAGTGAAGAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGACAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCCAAGACCAAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCAACCAAGACCATGGGGCATCAGGACAAAGAGATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
Db 889 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 949 CTACCCCTGCGCTGTGGGAGCTGTATCTGTAGAAAGCGAAGAGGCGTTGATGCGGATA 1008

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCCTTTAAGGTTCAAGAGCGCCTCAGGAAGAGCTCCAGGTGTGGCCCG 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGACGTTTGTGCAACCGGTTCTGTATACCTGGGCCAGGGATGAGGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGCGCTCTATGAGAACTAACAAAGAACAGGGGACCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
Db 1189 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCAAGTGGCGGAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTGTCAAGGGGACAGGAGGCCCTCCAGGCCACACAGCCCGGCTTCCCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
Db 1309 ATTGTACACTTCCCGGACAGGGACCTCTCTTCCACACCTGTCTGTGGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACACTACCTCTGTGGGACTCAGGGTGTGCCAGCCCGGCTGGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GGAGAACCTCAGGGTTCTTGAGATTCTTCAGTGAATCTCATGCCCCGCTCCCAAG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGACGTAGAGGGCCCTTGATTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1549 CGTGCATATCCAGAGGACAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGGTAGATCGGCTTTCGCTGACCCAGCACTGATTCGACCAAGGACACCCCTTCAGA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1669 GCTAGGAGACGAACAGCCGTGTGCTCCACCTCAGGGCTTGCCCTGCGGCTTCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1729 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755

RESULT 2
US-09-799-451-691
Sequence 691, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yuninging
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom

APPLICANT: Ghosh, Reena
APPLICANT: Dirmnac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_Fl_genes Version 2.0
SEQ ID NO 691
LENGTH: 2072
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1743)
US-09-799-451-691

Alignment Scores:
Pred. No.: 1.11e-199 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
DB: 4 Gaps: 1

US-10-071-838-2 (1-549) x US-09-799-451-691 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 163 ATGACGCTGTAGAGGTCCGGCGCAGTTGGTGGCCACAAGAGCGAGACATCATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAGAGGGGCTTAAGCCTTTTGA 282
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 283 AGCTACAAACAACAGTCGATCATTTGGGAGTGTACATGAGACGAGCTGCCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 343 ACTGCCGGGAGGCGAACAATTCCGGCGGAGATCAGCCGAAGAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 403 ATGCTGGAGACTGGGAGAAATACAAAAGCAGACGAAAGCTCATAGATGAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
DB 463 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 523 AAGTTGAAAAACCCCGAAGATACACAGATCATGAAGAGAGGCAAGGTCATCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 583 CACATCCAGCGCATGACCGGAGCGTAAGCGGCATTAAGAGCATATATCTTCAGG 642
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 643 GATCGATACGGAAACCAAGCAGCGGAATCTACATCTCTCTGGCATATGAGAGATAT 702
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
DB 703 AACCCGAGGTGGCTACTGACAGGGAACCTGAGCACAATCGCCGCTTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 763 CTTCCTGAGGAGATGATCTTGTGGGCACTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240

DB 823 CTGCAGGATTTCAACAGCCCAATATGCGGAGCCGTCCAGAGGGGCTCCAGAACCAACAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 883 CATGTGTAGCCAGCTCACAAACCAAGACCATGGGCATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300
DB 937 CTCACCTGCGCGCTGTGGGACGTGTATCTGTGAAGAGCGCAACAGCGTTGATGCCGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 997 ACAAGATCGCCTTAAAGTTACAGCAGAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1057 TGGGCACTTTTGTCAACCGGTTCTGTGATACCTGGCCAGGAGTGAAGACATGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1117 AAGCATCTTAGGGCTCTATGAAGAACTTAAACAAGAAAGCAGGGGGAACCTGCAACCCCA 1176
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1177 GCCAAACCGAGCAAGGCTGTGCGCATCAGGCTGTGCGGCTTCACTGCGGGAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1237 ACCCTGTCAAGGGGACAGGCAAGCCCTCCAGGCCACCAAGCCCGGTTCCCGCGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1297 ATTTGTCAGCTTCCCGCCAGGCAAGGCACTCTTCCACACCTGTCTGTGGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1357 GTCCGGGAAGACACTTACCTGTGGGCACTCAAGGTGTGCCAGCCCGGCTGTGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1417 GAGAGACCTCAGGGTCTCTGAGATCTTCAAGTGAACCTCAATGCCCGCTCCCAACG 1476
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1477 GACCTGACGTAGAGGGCCCTTGTTCCGCCATATGATTTCAACAGAGCTGTGGGTC 1536
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1537 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGCAAGGCTGAACACCTGCGGAG 1596
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1597 CGGGTGAGATCGGCTTTCTGCTGCAACCACTGATTCCGACCAAGGCAACCCCTTCAAG 1656
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1657 GCTAGGAGCAACAGCCTGTGTCTCCCACTCAGGGCCTTGTGCGGCTTCACTTGTG 1716
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 1717 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1743

RESULT 3
US-08-253-155A-9
Sequence 9, Application US/08253155A
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jeno

APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIT-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 8201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-253-155A-9

Alignment Scores:
Pred. No.: 1.35e-147 Length: 8201
Score: 2133.50 Matches: 408
Percent Similarity: 77.23% Conservative: 33
Best Local Similarity: 71.45% Mismatches: 58
Query Match: 71.52% Indels: 72
Gaps: 2

US-10-071-838-2 (1-549) x US-08-253-155A-9 (1-8201)

QY 1 MetaspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1697 ATGACATGCTAGAGATGACAGATAGTTTGACAGCACAGAGCGGAAAGACATATATG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTGAGCCCGTT--- 1813
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 1814 GGAATCAACAGACATGATCGTTTGGCATTTTGACATGACGAGGAGCTGCTCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 1874 ACTGACGGGAGCGGAGAAATTCGGCGGAGATGACAGAACGACGAACTGATGAGTGAAG 1933
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 1934 ATGCTGGAGAGATGGAGACATATTAAGCACAGTAGCAAACTCATATGATCGAGTGAAG 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGluMet 120
DB 1994 GGAATTCCTCATGAACATCCGGGGCCCGGTGTGTCTGCTCTGAAACATTCAGAAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 2054 AAGTTGAAAAACCCCGGAGATACCAAGATCATGAAGAGAGGGGCAAGAGGTTCATCTGAA 2113

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 2114 CACATTCACCAACATGACCTGAGCGTAGAGACGACTCTCCGAAACCATGTCTTCTTAGG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 2174 GATCGATATGAGCCCAAGCAGAGGAACTATTCTACATCTCTGCGCTATTCGAGTAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
DB 2234 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCACCGCTTGTTCTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 2294 CTGCTTAGAGAGACGCAATTCTGGCACTGGTGCAGCTGCTGCCAGTGAAGGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 2354 CTGCCAGGATTCACAGCCCAATGTGGACAGATCCAGGGGCTCCAAAGACCAACAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 2414 CATGTGTACCAAGTACACACCAACCAAGCATGTGCATCAGACCAAGAAAGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 2474 GGGCAGTGTGCTGTAGGCTGCTTCTCCGAACTGATGACGGATCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 2534 CTCACCTGCGCTGTGGAGCGTATTTGTGGAAGAGAAACAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 2594 ACCAGCATTTGCTTAAAGGTTCAGAGAGAGCGCTCATGAAAGACATCCAGGTGTGCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 2654 TGGGCACTGTGCGGAACCAATCTTGCATACCTGGCCATGAACGATGACACCGTGTCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 2714 AAGCATTTAGGGCTTTCGAAAGAACTAACAGAGCAAGGGGACCTGCCACCCCA 2773
QY 360 ----- 360
DB 2774 GGCCCAACAGCCCTGGGACGAAGGTGTGTGCAGAAACCCCCAGCCAGTCTGAACCTG 2833
QY 360 ----- 360
DB 2834 GGGGCAGTCCAGAGCAACCAATGCCCAAGGCTTCCCATGCCAGGACAGACAC 2893
QY 360 ----- 360
DB 2894 ACCCTCTCTGGGATCAGACACTACAGGCGTGTGTCACTGTCAGACCAAGGGGCC 2953
QY 361 ----- 370
DB 2954 ACACAGAGACCCCAAGACTCCAGAGATGCAAGCCMAACGCGAAGAGGTCTTGGCACC 3013
QY 370 rArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
DB 3014 CAGGCTGTGCGGCTTCACTGTGTGGAAGACCTCTGCAAGGGGTATAGGACGGCCCC 3073
QY 390 oProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410
DB 3074 TCAGGGCCACCAAGCTTCAGGCGGCCCATTTGCTCAGCTTCCCGCATGGGCATC 3133
QY 410 oArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyTh 430
DB 3134 TCGTTTTCACGCGCTCTGTGTGGGCTGTCCGGAGACACGTAACCTGTGGGCAC 3193
QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLe 450

Db 3194 TCAGGGTGTGCCAGCCCTGCGCTCAGGGAGCACTCAGGGTCTCGAGATTCT 3253
Qy 450 uGlnTrpAnsSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheAr 470
Db 3254 GGAATGGAAGTCAATGCCCCGGCTCCCAACGACCTGGATATAGGGGGCCCTGGTTCC 3313
Qy 470 gHisTyraSPheArgGlnSerCysTrpValArgAlaIleSerGlnGlnAspGlnLeuAl 490
Db 3314 CCATTATGATTTTGAACGAGCTGCTGGTCCGTCCATATCCAGAGAGACCAGCTGGC 3373
Qy 490 aProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 3374 CACCTGCTGGCAGCGCTGAACACTGCGGAGAG 3404
RESULT 4
US-09-620-312D-715
: Sequence 715, Application US/09620312D
: Patent No. 6569662
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyan
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aidong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yungqing
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhiwei
: APPLICANT: John, Tillinghast
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. 6569662el Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 784CIP2B
: CURRENT APPLICATION NUMBER: US/09/620,312D
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: NUMBER OF SEQ ID NOS: 1105
: SOFTWARE: pc_FL_genes Version 1.0
: SEQ ID NO 715
: LENGTH: 1981
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (59)..(1585)
US-09-620-312D-715
Alignment Scores:
Pred. No.: 3.3e-16 Length: 1981
Score: 344.00 Matches: 150
Percent Similarity: 40.18% Conservative: 77
Best Local Similarity: 26.55% Mismatches: 207
Query Match: 11.53% Indels: 131
DB: 4 Gaps: 26
US-10-071-838-2 (1-549) x US-09-620-312D-715 (1-1981)
Qy 8 GlySerTrpTrp-----AlaGlnGluArgGluAspIleIleMetLysTyrgLulys 24
Db 46 GGGCGCGGCTGGGATGCGCAAGACCAACGAGAGATGGCGCGCGCGCGCGCGCGCG 105
Qy 25 GlyHisArgAlaGlyLeuProGluAspLysGlyProLysPProPheArg-SerTyraSnaS 44

Db 106 GGAAGCCTGTGGGAAACCCGGGAGAGCCTGGCCCCAGGGCCCCGACCGCCCAACCACCA 165
Qy 44 naSnValAspHisIleuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgG1 64
Db 166 GGAATCAGCTCTGTGGGG-----TGTGACTCGGAGGCCAACGGCTTCGCCGAGCGCG 219
Qy 64 u-----AlaLysGlnI 68
Db 220 CATCGACAAGTTCGGCTTCATCGGCTCGGCTCGCAGGGCCGAGGGCGCGCTGGAGAACT 279
Qy 68 eaRgArgGluIle-----SerArgLysSerLysTrpValAspMetLeuGlyAspTrpG1 86
Db 280 ACCCTGGAAGTGTGAGAGCAGAGGAGTCCAAGTGGCTGACATGCTCAACAACCTGGGA 339
Qy 86 uLysTyra-----LysSerSerArgLysLeuIleAspArgAlaTyrlLysGlyMetProme 104
Db 340 CAAATGATGCGCAAGAACCAAAAAGATTGCTGCGGTCGCAAAAGGGCATCCCGCC 399
Qy 104 taSnIleArgGlyPromeTrpSerValLeuLeuAsnIleGluGluMetLysLeuLysAv 124
Db 400 TTCTTGGCGGGCGGCTGCTGGCAGTACCTGTCAAGAGCMAAGTGAAGTTACAGACAGAA 459
Qy 124 nProGlyArgTyrgIleMetLysGluLysGlyLysArgSerSerGluHisIleGlnAr 144
Db 460 CCCTGGAAGTTTGAAGAGCTG---GACATGTCCCTGGGAGACCCCAAGTGCGTGACGT 516
Qy 144 gIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrg1 164
Db 517 GATTGAGCGTGACCTGACCGGACGAGTTCCCATTCATGAATGTTGTGTCGCCGGGGGG 576
Qy 164 yThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrgLulGluTyraSnProGluVa 184
Db 577 CCACGGCCAGCAGACCTATTCCGTGTGTGAAGGCTTACACCGCTGTACCGGCCGAGGA 636
Qy 184 lGlyTyrcysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyrgLulGluProGluG1 204
Db 637 GGGCTACTGCGAGCGCCAGCGCCCATGTGCGCTGTCTGTCAATGATATGCTGCTGA 696
Qy 204 uAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPh 224
Db 697 GCAAGCCTTGTGCTGCTGTGATACAGATC-----TGTGAGAAATAC---CTGCCGGCTA 747
Qy 224 eHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnHisValValAl 244
Db 748 CTACAGCGAGAA-----CTGAGGCGCATCCAGCTGACGGGAGATCCTTTTCTC 798
Qy 244 aThrSerGln-----ProLysThrMetGlyHisGlnAspLysAspLeuCyseG1 261
Db 799 GCTGTTCAGAAAGTGTGCGCGGTGGCCACAGCACTCAGCCGTCAGAAAGATC----- 853
Qy 261 yGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLe 281
Db 854 -----GACCCGCTCTTATATAGACAGAAATGTTCATGTGCGCTTCCGGAACCTT 906
Qy 281 u-----ThrLeuArgLeuTrpAspValTyrlLeuValGluGlyGluGlnAlaLe 297
Db 907 GCCCTGAGCTGTGCTGCTGCTGCTGAGATGTTCTTCTGTGAAGG----- 955
Qy 297 uMetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerAr 317
Db 956 -----GTCAAGATCATCTCCGGGTGGGGCTG----- 982
Qy 317 gCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAs 337
Db 982 ----- 982
Qy 337 pThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLe 357
Db 983 -GTGCTGTGAAGCAGCGCTGGCTCCCTGAGAAAGTCAAAAGCTGCAAGGCCAGTA 1041
Qy 357 uPro-ProProAlaLysPProGlnGlnGlySerSerAlaSerArgProValProAlaSerA 377
Db 1042 CGAGACCATCGAGC-----GACTGCGGAGGCTCAGCCCCAAGATCA----- 1082

QY 377 rgGlyGlyLysThrLeuCysIysGlyAspArgGlnAlaProProGlyProProAlaArgP 397
Db 1083 -----TGcAGG----- 1088
QY 397 heProArgProIleTyrSer-----AlaSerPro-ProArgAlaProArg--- 411
Db 1089 -----AGGcCTTCTGTGTCAGAGGTGTGTGAGTTGCCGTGACAGAGCCcAGATTGA 1143
QY 412 SerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGln 431
Db 1144 GCGCGAACAACCTCATTCAGCTGCGCGCTGCGAGAGACCCG-----GGGTGAGCT 1194
QY 432 GlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrArgPheLeuGln 451
Db 1195 GCAGTGCCTGCTCCCGCCcAGGCT---GCAATGTGCAAGGCTATCTTGATGCAGAAC 1251
QY 452 TrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTyrPheArgHis 471
Db 1252 TGGTCC-----CCGGcCTGcCTTCAACCTTCACCATCCGCTGcCCCTAGA--- 1302
QY 472 TyrAspPheArgGlnSerCys-----TrpValArgAlaIleSerGlnGluAspGln 488
Db 1303 -----TGcCCcCTcCTcCTGCTCCAAAGCCAAAGCCcAGCCcCAA 1344
QY 489 LeuAlaProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAla 508
Db 1345 GCAGGcCCcAGAGAGAGACCGGAA---ACAGATGAAGGGAGAGAGGCAcCTGAGAAAGCC 1401
QY 509 ProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnPro----- 526
Db 1402 CCCAGcCCcCAATCAAGCCATGTTGTGTGCGCGCTGCAGAGAGATGCATGTCCcCCcACAGCA 1461
QY 527 CysAlaPro 529
Db 1462 TGTGcCCcC 1470

RESULT 5

US-09-774-528-336
Sequence 336, Application US/09774528
Patent No. 6743619

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6743619el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 336
LENGTH: 1823
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (136)..(1476)
US-09-774-528-336

Alignment Scores:

Pred. No.: 9.68e-16 Length: 1823
Score: 337.00 Matches: 157
Percent Similarity: 36.23% Conservative: 68
Best Local Similarity: 25.28% Mismatches: 231
Query Match: 11.30% Indels: 167
DB: 4 Gaps: 29

US-10-071-838-2 (1-549) x US-09-774-528-336 (1-1823)

QY 16 GluAspIleIleMetLysTyrGlu---LysGlnHisArgAlaGlyLeuProGluAsp--- 33
Db 154 GAGGACCTGTGcAGcCTCCCGAGcCTGcAGGATGACTCCAGcCTTGGGTCCGACTCA 213
QY 34 -----LysGlyProLysProPheArgSerTyrAsnAsnValAspHisLeuGlyIle 51
Db 214 GAGCTCAGCGGcCTGcCCCATATGcC-----CAGcCCGACcCTATGATTC 261
QY 52 ValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLysGlnIleArgGlu 71
Db 262 ATTGGGcGcAGCTCAGcAGcCCcAGGcCCGcCCcACCcCTGcAGcCTCATCCGCAA 321
QY 72 IleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGluLysTyrLysSer 91
Db 322 -----CGGAGATGAAGTGGGTGAAGATGACTCCGcACTGcGAGAAACCATGTCCCG 375
QY 92 Arg-----LysLeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyPro 109
Db 376 CGGTACAGAAAGTAAAGATGcAGTCCCGAAAGcCATCCGTGcCCCTGcCGcCCGA 435
QY 110 MetTrpSerValLeuLeuAsnIleGluGluMetLysLeuLysAsnProGlyArgTyrGln 129
Db 436 TGCTGcCCcCTGTGTGTGGGcCCcCATGTGTGcCAAGAAcAGcCCCTGcCACTATcAG 495
QY 130 IleMetLysGluLysGlyLysArgSerSerGlnHisIleGlnArgIleAspArgAspVal 149
Db 496 GAGCTGcCAGAG---GCCcCTGAGACCCcAGcTGAATGAGAACCATTTGcAGGAGcCTG 552
QY 150 SerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGlu 169
Db 553 CACcGTCAATTCCcCTGcCAGcAGcAGcTGTGTGcCTGcAGGcCCcAGcGcAGcGcG 612
QY 170 LeuLeuHisIleLeuLeuAlaTyrGlnGluTyrAsnProGluValGlyTyrCysArgAsp 189
Db 613 CTCCTGcAGGTGcTCAAGcCTTACcCTGTATGcACcGcGcAGcGcTACTGcCAGcGc 672
QY 190 LeuSerHisIleAlaIlePheLeuLeuTyrLeuProGluGluAspAlaPheTrpAla 209
Db 673 CAGGcGcCCcGTGcCTGTGcTGTATGcACcCTGcCCcCAGcAGcGcCTTCTGTGc 732
QY 210 LeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGly 229
Db 733 CTGcTGcAGATCTGT-----GAGcTCTACcCTCCcTGGTACTACGcCCcCCAC--- 780
QY 230 GlyThrValGlnGlyLeuGlnAspGlnGlnGlnHisValAlaIleThrSerGlnProLys 249
Db 781 -----ATGAGcCTGTGcCGcCTGcAGcCCcAGcGTTCATGcCCcCTGcCGc---CGG 831
QY 250 ThrMetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeu 269
Db 832 CTGcTTCGcAGcTGCACAGcACcCTGcAGcAGcTGGcCGcCTGcGACcCTGcTACTGc 891
QY 270 -----IleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeu 285
Db 892 CCCGAGTGTCTCTGTGcCTTTCGcCCcGTCCcCTTCCcCAcAGcTGTGcGTGc 951
QY 286 TrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIleThrArgIleAlaPhe 305
Db 952 TGGATGcCTTCTCCTCAGTGAAGGTGC-CAGcTACTGTTCG----- 992
QY 306 LysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCys 325
Db 993 -----TGTGcG----- 998

```
QY 326 AsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuYHisLeuArgAla 345
Db 999 -----GCTGACACTGGTGTGCGCTGGCGCTGGG-----CACTGCAGAGCA 1037
QY 346 SerMetLysLysLeuThrArgLysGlnGlyAspLeuPro---ProProAlaLysProGlu 364
Db 1038 GCGAGG-----GCGCTGCCCTGGCCTCTGGA-----GAC 1067
QY 365 GlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLysThrLeuCysLys 384
Db 1068 ACTGGGAGCCCTTCGAGCATCCCCCGCGCAGCTGCAGAGAGGCGCTTCATGTCACA 1127
QY 385 GlyAspArg-----GlnAlaProProGly 392
Db 1128 GGTGCACAGCGTGTGTCTGTCAGACGGGAGCTGCAGCGGAGATCAAGGCCAGCTGGC 1187
QY 393 ProProAlaArgPhe-----ProArgProIleTrpSerAlaSerProProArg 408
Db 1188 CCAGCTGCCGATTCGCGCGCGGAGACCCCGCCCG-----GCCACAGTTCGCTCGC 1241
QY 409 -----AlaProArgSerSerThr----- 414
Db 1242 CCGGGCCCAAGCATTTTGAGGCCCAAGCAGCTGCAGAGAGTGCAGAGGCGCAAGCC 1301
QY 415 -----ProCysProGlyGlyAla----- 420
Db 1302 TGAGGTGCTCGGATGTGTGTGACAGCCCCCGGAGAGCCAGACCAGCGCGGAAACC 1361
QY 421 -----ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeu 438
Db 1362 CCAGACCCCGCGCAAGAC---TTTCATGGGCTCTGACTCGGGCGCGGGCCCCC--- 1415
QY 439 AlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeu 458
Db 1416 CATCGAGGGGCCCCCGCCAGCCCAACGAGG-----CTCCACCTCTCTT 1457
QY 459 ProThrAspLeuAspValGlnGlyProTrp----- 468
Db 1458 CTTGACACACCCGCTTCGAGAGACCATGACTTAGTCCCCCAGTCTCAATTGCCTGA 1517
QY 469 -----PheArgHisTyrAspPheArgGlnSerCysTrp 479
Db 1518 TGGCTGATGCCAGCCCCGCAATATAGCACCCGCACTTACTCTTGGGACTCGGGGACTTGG 1577
QY 480 ValArgAlaIleSerGlnLysArgGlnLeuAlaProCysTrpGlnAlaGlnHisProAla 499
Db 1578 CTT-----CCTTCTGGCAAGGACCAAGCAGTGGG 1607
QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAsp----- 514
Db 1608 GAAGGAGG-AGTCTCTCGTGTACATAGTGGTCAAGCAGCTAGCATGAGAGGAGGTAC 1666
QY 515 -----GlnGlyThrProPheArgAlaArgAspGlnProCys 527
Db 1667 AGAGTGGGCACTGAGAGACCATGAACCGTCTGTGCTGCCAGGCGCTCAACA----- 1720
QY 528 AlaProThrSerGlyProCysLeuCysGlyLeuHisLeuGlnSerSerGlnPheProPro 547
Db 1721 GTACCAAGCCAGCACCAAGAGTCAAGGAGGGGTGGCTGAGTCAAGGAGACCCACA 1780
QY 548 Gly 548
Db 1781 GGG 1783
```

RESULT 6

```
US-09-620-312D-675
; Sequence 675, Application US/09620312D
```

; GENERAL INFORMATION:

```
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
```

```
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 675
; LENGTH: 3039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)..(2502)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3039)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-675
```

Alignment Scores:

Pred. No.:	1,36e-09	Length:	3039
Score:	258.00	Matches:	157
Percent Similarity:	35.89%	Conservative:	77
Best Local Similarity:	24.08%	Mismatches:	236
Query Match:	8.65%	Indels:	183
DB:	4	Gaps:	30

```
US-10-071-838-2 (1-549) x US-09-620-312D-675 (1-3039)
```

```
QY 9 SerTrpTrpAlaGlnGluArgGluAspIleIleMetLysTyr---GluLysGlnHisArg 27
Db 220 AGCATATGCCCCCAGAG-----ATCTGGCCAAAGTACACGACAGAGAGAGTCA 270
QY 28 AlaGlyLeuProGlu-----AspLysGly 35
Db 271 GCAAGACCAACCAAGATTCTACTACGATGAGTTGGTTCCCGTGTGTAACAAGAGAAGGT 330
QY 36 ProLysProPheArgSerTyrAsnAsnValAsp----- 47
Db 331 GATGAGCCTGGCTCCAGTCTGTGCGCAACTCCCTGTGATGAGAGATGCTCCACAGAGG 390
QY 48 -----HisLeuGlyIleValHisGluThrGluLeuProProLeuThrAla 62
Db 391 CTGCGGTGCAGGCCCACTGAGTTCAACCATTAACACGATGGGGGATCTCACC--- 447
QY 63 ArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAspMetLeu 82
Db 447 ----- 447
QY 83 GlyAspTrpGluLys-----TyrLysSerSerArgLysLeuIleAspArgAla 98
Db 448 -----TGGACACAGATTGCCGTCTCCCTAACCCCGCTGTGAGAGCTCCGCTGTTG 501
QY 99 TyrLysGlyMetPrometAsnIleArgGlyPrometTrpSerValLeuLeuAsnIleGlu 118
Db 99 TyrLysGlyMetPrometAsnIleArgGlyPrometTrpSerValLeuLeuAsnIleGlu 118
```

Db 502 CTGGCCGCGATCCACATGCGATGAGGCCACAGCTGTGATGCGGCTCTCTGGGCGCTG 561
 QY 119 GlnMetLysLeuLysAsnProGlyArgTyr---GlnLeuMetLysGlyLysArg 137
 Db 562 CAGAGAGAGAGAACTGTAGCTGTCTCTACCGCGAGATTGTGAAGAACAGCTCCACAGAT 621
 QY 138 SerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIle 157
 Db 622 GAGACCATCGCTGCCAAGCAGATGAGAGAGACCTGCTCCGACCCATGCCCCAGAACGCC 681
 QY 158 PhePheArgAspArgTyrGlyThrLysGlnArgGlnLeuHisIleLeuLeuAlaTyr 177
 Db 682 TGCTTCGCCAGCATGGTAGCATCGGGGTGCCCGCCCTGCGCAGGGTGTCCGGGCGCTG 741
 QY 178 GlnGlnTyrAsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPhe 197
 Db 742 GCCTGGCTCTTACCCAGAGATCGGCTACTGCGAGGGCACCAGGATGTGCGCGCTGCTC 801
 QY 198 LeuLeuTyrLeuProGlnGlnAspAlaPheTrp-----AlaLeuValGln---Leu 213
 Db 802 CTGCTGTCTCTGAGAGAGAGAGAGCGCTTCTGATGATGCTCTGCCATCATCGAGAGACCTG 861
 QY 214 LeuAlaSerGlnArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGln 233
 Db 862 CTCCCGCCCTCTACTTCAGCACC-----ACCCCTGCTG 894
 QY 234 GlyLeuGlnAspGlnGln-----GlnHisValAlaAlaThrSerGlnProLysThr 250
 Db 895 GGTTCCAGACTGACCCAGCGGCTCTGCGCCACCTCATTTGTCCAGTAACCTGCTCGCTG 954
 QY 251 MetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIle 270
 Db 955 GACAAGCTGCTCCAGAGCATGACATTGAGCTGTCCGTATGATCACTGACACTGCTGCTC 1014
 QY 271 ArgIleLeuIleAspGlyLysSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeu 290
 Db 1015 ACGGCTTCGCCAGCGGTGTGACATCATGCTCTCTGCGCATCTGGGACTGTTTTC 1074
 QY 291 ValGlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnLys 310
 Db 1075 TACGAGGGCTCCCGGCTGTGTTCAGCTACGCTGGGATGCTGACCTCAAGAGAGAA 1134
 QY 311 ArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAsp 330
 Db 1135 GAGCTGATCCAGTCAGAGAACTCG-----GCCCTCATCTTCAACACGCTATGAGAT 1185
 QY 331 ThrTrpAlaArgAspGlnAspThr-----ValLeuLysHisLeuArgAla 345
 Db 1186 ATCCCGTCGAGATGAGAGCGGAGCTGCTTCTGGGGTGGCCATGCGGCTGCGCGGC 1245
 QY 346 SerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro----- 359
 Db 1246 TCCCTCAC-CGATGTGGCGT-----GGAGACTCAGCGCGCAAGCTCTGCGCTATCT 1298
 QY 360 -----ProAlaLysProGlnGlnLysSer 367
 Db 1299 CATTCAGACACGAGGCGAGCTCTCTGGGGCGCGCACCTCTCACCACTCTCTCAGGTTGT 1358
 QY 368 SerAlaSerArgProVal-----ProAlaSerArgLysGly--- 379
 Db 1359 TCGCCCGCAGACCCAGGAGAGATCCACCATCATGCTCTGCTCTTCGGGGAGGATGA 1418
 QY 380 -----LysThrLeuCysLysGlyAspArgGlnAlaProProGly 392
 Db 1419 CCTGAGGCACTCAAGGCCAAGAACATCAAGCAGACGGAACGTGTGCTGACCTCGGGA 1478
 QY 393 Pro---ProAlaArg---PheProArgProIleTrpSerAlaSerProProArgAlaPro 410
 Db 1479 AGCCATCTGCGCGTGCACGCCACTTCCAGTg-----CACAGACCCCAAAATGTCAG 1532
 QY 411 ArgSerSerThrProCysProGlyGlyAlaValArgGlnAspThrTyrProValGlyThr 430
 Db 1533 CGTGTGAGTGCAGCTCCCTGCGTGT-----ACC 1565

QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGln-GlySerTrpAspPheLe 450
 Db 1566 AAACACGGCCCTTACTCTCTCAACCCCTTGGTGGGCTGTGTTCACCTGTGGCAGAGCT 1625
 QY 450 uGlnTrpAsn---SerMetProArgLeuProThrAsp----- 461
 Db 1626 GACTCCAGACTATAGCATGAGAGCCACGAGCGGACCAAGAACTACGTGCGGTGCTC 1685
 QY 462 -----LeuAspValGlnGlyProTrpPheArgHis 471
 Db 1686 ACCGACCCACCGCGCGCGAGCCAGGCCCTGCTGACTTTGAG-----CGGCA 1733
 QY 471 sTyrAsp-----PheArgGlnSerCysTrpValArgAlaIleSerGlnGlnAs 487
 Db 1734 CGACGACGACGAGCTGGGCTTCCGCAAGAACAGACATCATCAATGCTGTCTCAGAAAGA 1793
 QY 487 pGlnLeuAlaProCysTrpGlnAlaGlu-----HisProAlaG1 500
 Db 1794 CGAG-----CACTGCTGGGTGGGGAGCTCAACGGCTGCGAGCTGTTTCCAGCCAA 1847
 QY 500 uArgVal-ArgSer---AlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProp 519
 Db 1848 GTTCGTGAGAGTCTGTGATGACCGCAGCAAGAGTACTCCATCCGCGGGGATGACTCGGT 1907
 QY 519 heArg-----AlaArgAspGlnGlnProCysa 528
 Db 1908 GACGAGGGGGGTACAGACCTCGTGGCAGGAGACCTGTGCGCGCCCTTAAGGCCCTGTT 1967
 QY 528 lapro-----ThrSerGlyProCysLeu 535
 Db 1968 CGAACATGACTGAAGAACCATCCCTGCTTG 1999

RESULT 7
 US-09-620-312D-676
 ; Sequence 676, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Wang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Wang, Duntui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE REFERENCE: 784CIP28
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pc_FL_genes Version 1.0
 ; SEQ ID NO 676
 ; LENGTH: 2955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (169)..(2418)

FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2955)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-676

Alignment Scores:
Pred. No.: 3.06e-09 Length: 2955
Score: 253.00 Matches: 156
Percent Similarity: 35.80% Conservative: 71
Best Local Similarity: 24.61% Mismatches: 233
Query Match: 8.48% Indels: 175
DB: 4 Gaps: 28

US-10-071-838-2 (1-549) x US-09-620-312D-676 (1-2955)

QY 9 SerTrpTrpAlaGlnGluArgGluAspIleIleMetLysTyr--GluLysGlnHisArg 27
DB 220 AGCATATGCCCCAGAG-----ATCTTGCCCAAGTACACGAGAGGAAGAGTCA 270
QY 28 AlaGlyLeuProGlu-----AspLysGly 35
DB 271 GCAGAGCAACCAAGATTCTACTACGATGATTGGTTTCCGTGTGTACAGGAAGAAGT 330
QY 36 ProLysProPheArgSerTyrAsnAsnValAsp----- 47
DB 331 GATGAGCCTGGCTCCAGTCTGCTGGCGAAGTCCCTCTGATGAGGATGCTCCACAGAG 390
QY 48 -----HisLeuGlyIleValHisGluThrGluLeuProProLeuThrAla 62
DB 391 CTGGCGGTGGCAGGCCACCTGGAAGTTCACCCATAACGATGTGGGGATCTCACC--- 447
QY 63 ArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAspMetLeu 82
DB 447 ----- 447
QY 83 GlyAspTrpGluLys-----TyrLysSerSerArgLysLeuIleAspArgAla 98
DB 448 -----TGGACAAAGATTGCCGTCTCCCTACCCCGCTCTGAGAAGCTCCGCTCCGTGTG 501
QY 99 TyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlu 118
DB 502 CTGGCCCGGCATCCACATGGCATGAGGCCACAGCTGTGATGCGGCTCTGGGGCCCTG 561
QY 119 GluMetLysLeuLysAsnProGlyArgTyr--GlnIleMetLysGluLysGlyLysArg 137
DB 562 CAGAGAAGAGGAAGTCTGAGCTGTCTTACCGCGAATTGTGAAGAAGACAGCTCCAAGCAT 621
QY 138 SerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIle 157
DB 622 GAGACCATCGCTGCCAAGCATGAGAGAGACCTGCTCCGACCATGCGCCAGCAAGCC 681
QY 158 PhePheArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyr 177
DB 682 TGCTTCCGACGATGGGTAGCATCGGGGTGCCCCGCTGCGAGAGTGTCCGGGCCCTG 741
QY 178 GluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPhe 197
DB 742 GCCTGGCTTACCCAGAGATCGGCTACTGCCAGGCAACCGCATGCTGGCCCTGCTC 801
QY 198 LeuLeuTyrLeuProGlnGluAspAlaPheTyr-----AlaLeuValGln--Leu 213
DB 802 CTGCTGTCTCTGAGGAGAGAGAGCGCTTCTGATGATGTCTGCATCATCGAGGACCTG 861
QY 214 LeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGln 233
DB 862 CTCCCGCCCTCTACTTCAAGACC-----ACCTGCTG 894
QY 234 GlyLeuGlnAspGlnGln-----GlnHisValAlaAlaThrSerGlnProLysThr 250
DB 895 GGTGTCCAGACTGACCAAGCGGATCTGCGCACTCATTTGTCAGTACTGCTGCGCTG 954
QY 251 MetGlnHisGlnAspLysValAspLeuCysGlyGlnCysSerProLeuGlyCysValLeuIle 270

DB 955 GACAAGCTGCTCCAGAGCATGACATTTGAGCTGTCCCTGATCACACTGCACTGTCTC 1014
QY 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeu 290
DB 1015 ACGGCTTCGCCAGCGGTGGTGGACATCAAGCTGTCTCTGCGCATCTGGGACCTGTTTTC 1074
QY 291 ValGluGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLys 310
DB 1075 TACGAGGGCTCCCGGTGCTGTTCAGCTCAAGCTGGCGATGCTGCACCTCAAGAGAGAA 1134
QY 311 ArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAsp 330
DB 1135 GAGCTGATCCAGTCAAGAACTCG-----GCCTCATCTTCAACACGCTATCGGAT 1185
QY 331 ThrTrpAlaArgAspGluAspThr-----ValLeuLysHisLeuArgAla 345
DB 1186 ATCCCGTCGAGATGAGAGACCGGAGCTGCTTGGGGGTGGCCATGGCGCTGGCCGC 1245
QY 346 SerMetLysLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGln 365
DB 1246 TCCCTCAC-CGATGTGCGCT-----GGAGACTCAGCGCCGCAAGCT-----CCT 1289
QY 366 GlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGly 385
DB 1290 GGCCTATCTCATTTGACAGACCAAGGCCAGCT----- 1319
QY 386 AspArgGlnAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSer 405
DB 1320 -----CCTGGGGCGGCACCCCTCACCAACCTTCTCAGGTGT 1358
QY 406 ProProArgAlaProArgSerSerThrProCysProGlyGlyAlaValArgGluAspThr 425
DB 1359 TGGCCGACAGACCCAGGAGAGAGTCAACCATCACTGCTCTGCTCTTCCG--GGAGA 1415
QY 426 TyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyGly--ProGln 444
DB 1416 TGACCTGAGGCACTCAAGGCCAAGACATCAAGCAAGACGGAAGTGGCTGACCTCCG 1475
QY 445 GlySerTrp-----ArgPheLeuGlnTrp 452
DB 1476 GGAAGCCATCTCGCGGTGGACAGCCACTTCCAGTGCACAGACCCCAAACTGACGCGT 1535
QY 453 AsnSerMetProArgLeu-----ProThr-Asp----- 461
DB 1536 GGAAGTGAATCAAGTATAGCATGAGAGGCCACAGCGGAGCAAGAACTAGCTGGC 1595
QY 462 -----LeuAspValGluGlyProTrpPh 469
DB 1596 GTGCTCAGCAGCAACCGCGCGGAGCCAGGAGCCCTGTGACTTTGAG----- 1644
QY 469 eArgHisTyrAsp-----PheArgGlnSerCysTrpValAlaArgAlaIleSerGly 485
DB 1645 -CGGACAGACGACGAGAGCTGGGCTTCCGCAAGAACGACATCATCAATCGTCTCA 1703
QY 485 nGluAspGlnLeuAlaProCysTrpGlnAlaGlu-----HisPr 498
DB 1704 GAAGGACGAG-----CACTGCTGGGTGGGAGACTCAACGGCTGCGAGGCTGTTCC 1757
QY 498 AlaGluArgVal-ArgSer--AlaPheAlaAlaProSerThrAspSerAspGlnGlyT 517
DB 1758 AGCCAAAGTTGTTGAAGTCTGTATGAGCGCAGCAAGAGTACTTCATCGCGGGGATGA 1817
QY 517 hPProPheArg-----AlaArgAspGluGlnP 526
DB 1818 CTGGGTGACGAGGGGTCAAGACCTGTGCGAGGAGGACCTGTGCGCGCTTAAGGC 1877
QY 526 roCysAlaPro-----ThrSerGlyProCysLeu 535
DB 1878 CTTGTTGAAATGAGTGAAGAAGCAATCCCTGCTTG 1915
RESULT 8
US-09-248-796A-5457

```
/ Sequence 5457, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 5457
/ LENGTH: 1083
/ TYPE: DNA
/ ORGANISM: Candida albicans
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (37)
/ OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-5457
```

```
Alignment Scores:
Pred. No.: 9.71e-09 Length: 1083
Score: 238.00 Matches: 72
Percent Similarity: 47.12% Conservative: 59
Best Local Similarity: 25.90% Mismatches: 127
Query Match: 7.98% Indels: 20
DB: 4 Gaps: 10
```

```
US-10-071-838-2 (1-549) x US-09-248-796A-5457 (1-1083)
```

```
QY 64 GUAAlaLysGlnIleArgArgGluIleSerArgLysSerLys---TrpValAspMetLeu 82
DB ::::: 13 GACGCCGCATTTGGGGAACACGAANTGCCAACAAATTTGAATAATGTGTTGATTATTATT 72
QY 83 ---GlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys--- 100
DB 73 CAACCGATGACGAATAATTATCTTTGTTGAATAAACCCTTTA-----TTTATAAACTT 126
QY 101 -----GlyMetPrometAsnIleArgGlyPrometTrpSerValLeuLeuAsnIle 117
DB 127 ATTCGAGTAGGATTACCAATAGATTAAAGAGAGAAATATGGGAATTGACTTGGGTGC 186
QY 118 GluGluMetLysLeuLysAsnProGlyArgTyr---GlnIleMetLysGluLysGlyLys 136
DB 187 ATGATTATTACGATTAGAACACACGAGGTGAATATTCTCAACTATTAGAAAGAACATATAAGAC 246
QY 137 ArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHis 156
DB ::::: 247 AAGAAATCGTTTGTCTATTGAAGAGATTGAATAAGATTGAATAGATCATTAACCGAGTAT 306
QY 157 IlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuAla 176
DB ::::: 307 GCTGCATATCAATCACTGAGGGAATTGACAGG-----TTAAGAAGATATTGACAGCC 360
QY 177 TyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeu 196
DB 361 TATCTTGTGAAAATCTCGATGTGTTATTCTCAAGCTATGATATAGTGTGTCAGCA 420
QY 197 PheLeuLeuTyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuAlaSer 216
DB 421 TTACTTATATACATGCTGAAGACAAACCATTTTGGCGGTGAATGTTTATGT----- 474
QY 217 GluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly---Leu 235
DB 475 ---GACAGATAAGTACCTGGGTATTATTCAAAA-----ACCAATGATGAACCTTTA 522
QY 236 GlnAspGlnGln-----GlnHisValValAlaThrSerGlnProLysThrMetGlyHis 253
DB 523 TTGGACCAAAAGTATTGAACTTTGGTTCAAGATATCCATGCTTATGCTTTGGGAACAT 582
```

```
QY 254 GlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeu 273
DB 583 ATCACCAAAAATGATATATCAATATATCAGTGTTCATTACCGGTGCTTTTATCATATATAC 642
QY 274 IleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGly 293
DB 643 TTGCTTCAATGCCATTAGTTTTCATTCGCCAATCCTTGATATATTTTTCATCAAGCA 702
QY 294 GluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThr 313
DB 703 CCAAAAACCTTATTTCAGTGTTCATTCGCTATTTTAAAGCAAAATGAGAGAACTTTTG 762
QY 314 LysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThr 331
DB 763 CAGACTGAAGATGATGAACATTCATATCTATATTAAGATTAATTTTCATTA 816
```

RESULT 9

```
US-09-270-767-12482
/ Sequence 12482, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 12482
/ LENGTH: 1887
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-12482
```

Alignment Scores:

```
Pred. No.: 3.2e-07 Length: 1887
Score: 222.00 Matches: 129
Percent Similarity: 36.62% Conservative: 64
Best Local Similarity: 24.48% Mismatches: 198
Query Match: 7.44% Indels: 138
DB: 4 Gaps: 23
```

```
US-10-071-838-2 (1-549) x US-09-270-767-12482 (1-1887)
```

```
QY 47 AspHisLeuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLys 66
DB 105 GACCTTCTGCTCGAGTTCAAT-----GTTCTTTGAGTCGCAACGCCGCTTAAG 152
QY 67 -----GlnIleArg 69
DB 153 TACGACATTTCTGAGACAGACAGACCGCTCTGATGAACACATTCAAGACACAGTTCAGC 212
QY 70 ArgGluIleSerArgLysSerLysTrpValAspMetLeuGlyAspTyrGluLysTyr--- 88
DB 213 GCAGAGATCATACAGAACAGAGAG-----GAAAAGATGTCGCTGGAGCGCATTTTC 266
QY 89 -----LysSerSerArgLysLeuIleAspArgAlaTyr 99
DB 267 CGGACTTCGACCGGGCATTTGAATCTCCGACACAGCAGCATGATTAATTAATAGTG 326
QY 100 LysGlyMetPrometAsnIleArgGlyPrometTrpSerValLeuLeuAsnIleGlnGlu 119
DB 327 GAGGGTATTCCTGACAAACTGCGCAGAGATCTGCTGATATTTCTGGGGGCAATCCAC 386
QY 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGly-----Lys 136
DB 387 GACAAAGAGATGAATCCCGGAGCTGACGAGATCTCTAGAGAAAGCGGCTTGACATAAG 446
QY 137 ArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHis 156
DB 447 AACTGCTTTGCCAC---GATGAATTTGATCGGACTTGCCGCTTCGCTGCCGAGCAT 503
```

```
QY 157 IlaPhePheArg-----AsparGTYrGlyThrLySGlnArgGluLeuLeuHisIleLeu 174
Db 504 CCGGCATTTCAAAGCACCAGTCATAGTGTCTGCGGAGA-----GTCCTT 551
QY 175 LeuAlaTYrGluGluTYrAsnProGluValGlyTYrCysArgAspLeuSerHisIleAla 194
Db 552 CAGGCCCTATGCGGTGCGGAATCCACAAGTGGGCTACTGTACGGCCATGATATGTGTGCG 611
QY 195 AlaLeuPheLeuLeuTYrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeu 214
Db 612 TCGGTGTTCTGCTCTTCTGTGATGAGGAGAACGCCATTCTGATGTGCGCCACCTGTGC 671
QY 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 234
Db 672 -----GAGAACTTAATAACCGACTACTACAGATAAA-----GTGGTGGGC 713
QY 235 LeuGlnAspGlnGln-----GluHisValAlaAlaThrSerGlnProLysThrMet 251
Db 714 GCCCAATCGATCAGGGTGTGCTTAATGAGCTAGTGAGACGCATTGCGCTGATTGAC 773
QY 252 GlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysValIleArg 271
Db 774 GGCCATCTGAGCAGTGGGTGTGATTAAGATGATTTCC---ATTTCCTGGTTCCTTACC 830
QY 272 IleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTYrLeuVal 291
Db 831 ATCTTCATGAGCGTGATCAGCTACGAGAGCTCGCTGCACATCCTGACTGCTTTTTCAC 890
QY 292 GluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArg 311
Db 891 GAGGGCGCCAAAGATCATTTTATGATTTCACTGTCAGATCATTGAATGGAACAGGACAAA 950
QY 312 LeuThrLysThrSerArgCysGlyProTrpAlaArg-PheCysAsnArgPheValAspThr 331
Db 951 CTGTGATCTGCCAAGATGATGGCGAAGCCATGTAGTTTGCAGA-----ACTAC 1001
QY 331 TrpAlaArgAspGluAspThrValLeu----- 340
Db 1002 CTGGAGG-----GAGTTTACAATCCAGAGTACGAGTGCTCTCTACGACGGAACAACGG 1055
QY 341 -----LysHisLysValArgAlaSerMetLysLysLeuThrArg 352
Db 1056 AAGATGAGCGCTAAAGTCCAAACAAACAACAGTTCAGACCTGATCCATGAGGCATACACA 1115
QY 352 GlyValGlnGlyAspLeuProProAlaLysProGluGlnGlySerSerAlaSerArgPr 372
Db 1116 AAAAGATTGGAT-----GAGGAGTATCACACAGCAGCGCATCGAGGAG 1160
QY 372 ovalProAlaSerArgGlyGly----- 379
Db 1161 TTACGACAACAGCATCGCGGTTAACGATGCGACAGTTCGATATCGACAAGAAAGACC 1220
QY 380 -----LysThr 381
Db 1221 ATTGTCAAGGCTTACGTCCAGAAATCCGATTTTCAACCGCACTGAACCTGATATGCTGCTC 1280
QY 381 rLeuCysLysGlyAspArgGlnAlaProProGlyPro----- 393
Db 1281 ACTATCATCCGGAGGAAACACACGCCCTTAATCCCTGCAACAGCAGCAGAGAAGGTT 1340
QY 394 -ProAlaArgPheProArg---ProIleTrpSerAlaSerProProArgAlaProArgSe 412
Db 1341 CAATGCCCGTGTGTGAGACTCCCAAGCTGTCTCTCAGT--CGACAAGCAGACCGATC 1397
QY 412 rSerThrProCysProGlyGlyAlaValAlaArgGluAspThrTYrProValGlyThrGlnG 432
Db 1398 CAAGATGCGGCGCTCC-GGAGGC--AGATACGAGGCTTACAGTGTCACTATGA-GG 1452
QY 432 yValPro-----SerProAlaLeuAlaGlnGlyGlyProGlnGlySerTx 447
Db 1453 TATTCACACACACTTTTACGGAACCTGACACCCCTGGCGTAAAGTGCCTACGCGTGACATTG 1512
QY 447 parGpheLeu-----GlnTrpAsnSerMetProArgLeuProThrAs 461
```

```
Db 1513 GCGAAAATTTGTTAGGCTTACTGATAGAGGAACTGTGTCTCGACTTCGACAGC 1572
QY 461 pleuAspValGluGlyPro 467
Db 1573 TCATCAATGATGATGGGCTT 1591

RESULT 10
US-09-270-767-9899/c
; Sequence 9899, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9899
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9899

Alignment Scores:
Pred. No.: 1.05e-05 Length: 806
Score: 194.50 Matches: 79
Percent Similarity: 41.14% Conservative: 44
Best Local Similarity: 26.42% Mismatches: 111
Query Match: 6.52% Indels: 67
DB: 4 Gaps: 9

US-10-071-838-2 (1-549) x US-09-270-767-9899 (1-806)
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 788 GCGGCTTCAGAAAGCACTGCGGTGAAGATCTGGCAGAACTAGCCAACTGAGGGC--- 732
QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGlyLysGlyLysArgSerSerGlu 140
Db 731 AGAATGGAATGAAACGACAAAGTACAGATCTTAATCACCAAAGAAACCAATGTGAGACC 672
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 671 GTTATCCAG-----CGGACATCCATCGCACTTCCCGCGCAAAATGCTTCAAA 621
QY 161 AspArgTYrGlyThrLySGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 620 GAAATTTGGCGGTTCTGGCCAAAGATGCACCTTTTAAAGGTGCCAAAGCGTATGCCGTTCA 561
QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTYr 200
Db 560 GACAGCGAGGTGGATATTGTCAAGGCTTAAGTTTCATAGCAGCTAGTCTCTCAT 501
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 500 ATGCTGAGGAGGATGCTTCTGTGTTCTGTAGCGCTTATATGACGAC----- 453
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAsp----- 237
Db 452 -----TATGGGCTCCGATCTCTACAAA 429
QY 238 -----GlnGlnGluHisValAlaAlaThrSer 246
Db 428 GCGGATTCGAGGTCTTACCTTCTGTTACCAACTGAGCGAGCTGATCAAGATCAG 369
QY 247 GlnProLysThrMetGlyHis-GlnAspLysLysAspLeuCysGlyGlnCysSerProLe 266
Db 368 CTGCCCCAAGCTGCACGAACACTTACAGCGCTGCGGCATTGAGAGCAGCATGTAAGCTCC 309
QY 266 uGlyCysValIleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTr 286
```

Db 308 CAATGG-TTCTTAACCTTATATACAGCGGATTCCTCATTTGTGTTTCACGTCCT 250
Qy 286 pAspValTyrLeuValGluGluGlnAlaLeuMetProIleThrArgIleAlaPheLy 306
Db 249 GGATGTGTTCTACTGACGGA-----CTACTGTG----- 219
Qy 306 sValGlnGlnYsArgLeuThrLysThrSerArgCysGlyProTrrpAlaArgPheCysAs 326
Db 218 -CTCTCCAGGTGGCCGTACTCTCTTATCAATCTGTGAATCC----- 177
Qy 326 nArgPheValAspThrTrrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSe 346
Db 176 -----GACTTGGCGCACTCGATTGTGAGGGCATTCTAAAGTAATTCCGGGTAAAC 127
Qy 346 rMetLysLysLeuThrArgLysGlnGlyAspLeuProProAlaLysProGluGlnG 366
Db 126 GTTCCGCAAAAAGTGCCCGCAGCTCC-----AGTCAGGC-ACGCAAG 86
Qy 366 ySerSerAlaSer-----ArgProValProAlaSerArgGly 379
Db 85 TGATGAAGCAAGCCTCGCAACGTAGATAAAAAAAGTGAAGCAGTACGAGAGAGA 31

RESULT 11

US-08-363-300-1
Sequence 1, Application US/08363300
Patent No. 5700927

GENERAL INFORMATION:

APPLICANT: Zon, Leonard and Richardson, Paul
TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,300
FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04590/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4039 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 403..3829
US-08-363-300-1

Alignment Scores:

Pred. No.: 0.000768
Score: 182.50
Percent Similarity: 34.59%
Best Local Similarity: 22.52%
Query Match: 6.12%
DB: 1
Length: 4039
Matches: 125
Conservative: 67
Mismatches: 227
Indels: 139
Gaps: 18

US-10-071-838-2 (1-549) x US-08-363-300-1 (1-4039)
Qy 13 GlnGluArgGluAspIleIle-----MetLysTyrGluLysGluYHisArgAlaGly 29
Db 2408 AGAAAAGAGAGACGTGACCGAGCTTCGAGAGCTGTGAAAAAGCCATCTTGACGA 2467
Qy 30 LeuProGluAspLysGly---ProLysProPheArgSerTyrAsnAsnValAspHis 48
Db 2468 GATCCTGCCTCGTCAGGATGAGAGAGAGAAATCAGAGCTACAG-----CCT 2515
Qy 49 LeuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLysGlnIle 68
Db 2516 CTGAAAACGATTGTGTGAACAAACGCTCAAGCTTGACTATGAAGAAATCCTCCGTGC 2575
Qy 69 ArgArgGluIleSerArgLysSerLysTrrpValAspMetLeuGly----- 83
Db 2576 TT-AAAGAGTCACT-----ACAGTGTGGGAAAAGATGCTTAAGCACTCCAGGAAGATCC 2628
Qy 84 -----AspTrrpGluLysTrrpLysSerSerArgLysLeuIleAspArgAlaTyr 99
Db 2629 AAAATTAGTTTGACATGAAAAAAGTCACTACGT-----GTTGGG 2670
Qy 100 LysGlyMetProMetAsnIleArgGlyProMetTrrpSerValLeuLeuAsnIleGluGlu 119
Db 2671 CAAGGTGTGCCACGTATCACCAGCGAGTGAATCTGGAAATTCTAGCTGACGAGTTCAC 2730
Qy 120 MetLysLeuLysAsnProGlyArgTyrGln-----IleMetLysGluLysGly 135
Db 2731 CTAAACACCCATTCTCTAGTAACACAGACCCAAAGACGTCCTTCAAAAGAGCTCCTG 2790
Qy 136 LysArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLys 155
Db 2791 AAGAGCTGACCTCGCAGCAGCAGCCATTCTCATTCGACCTCGGGCGAACCTTCCAACA 2850
Qy 156 HisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeuHisIleLeu 175
Db 2851 CATCCATACTTCTCTGCCAGCTTGGACGAGTCACTGTCACTTTAACAACATTCTGAAG 2910
Qy 176 AlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAla 195
Db 2911 GCCTACTCGCTTGTGACACAGAGGTTGGATTACTGCCAAGGTCTCAGCTTGTGGCAGGC 2970
Qy 196 LeuPheLeuLeuTyrLeuProGluGluAspAlaPheTrrpAlaLeuValGlnLeuAla 215
Db 2971 ATTTGCTTCTTCAACATGAGTAGGAGAGAGCGCTTCAAGATGCTCAAGTCTGTATGTTT 3030
Qy 216 SerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeu 235
Db 3031 GAC-----ATGGGCTGCCGGAACAGTATCGGCCAGACATGATTATTTGGCAGATCCAG 3084
Qy 236 GlnAspGlnGlnGlnHisValAlaIleThrSerGlnProLysThrMetGlyHisGlnAsp 255
Db 3085 ATGTACACAGCTGTACAGGCTCTCCACGATTACCACCGAGACCTTACAACCACTTGAA 3144
Qy 256 LysLysAspLeuLysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAsp 275
Db 3145 GAGCAGAGAGACTGGCCCCCTACGTACGGGCTCGTGTCTCACCCGTGTCGCTCA 3204
Qy 276 GlyIleSerLeuGlyLeuThrLeuArgLeuTrrpAspValTyrLeuValGluGlyGln 295
Db 3205 CAGTCCCACTCGCTTGTAGCCAGAGCTTTGATATGATCTTCTTCAAGGATCAGAG 3264
Qy 296 AlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThr 315
Db 3265 -----GTCAATATTAAAGTAAAGCTTAAAGCTTTTG----- 3294
Qy 316 SerArgCysGlyProTrrpAlaArgPheCysAsnArgPheValAspThrTrrpAlaArgAsp 335
Db 3295 -----GGAGCCAT 3303
Qy 336 GluAspThrValLeuLysHis-----LeuArgAlaSerMetLysLysLeuThrArgLys 353
Db 3304 AAGCCCTTGATTCTACAGCATGAGAACCTGAAACCATCGTGAAGCTTCATAAAGAACACA 3363


```
QY      354 GInGlyAspLeuProProAlaLysProGInGlySerSerAlaSerArgProVal 373
      3364 CTCCCCAACCTGGG-CTGTGTGACAGATGAGAGAACCATCACTCAGGT----- 3410
QY      374 ProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaProProGlyPro 393
      3411 -----GTTTGAATGACATCGCCCAAGCAGCTCCAGGCTTA 3446
QY      394 ProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSerSer 413
      3447 TGAGGTCGAGTACCACGTCGT----- 3467
QY      414 ThrProCysProGlyGlyAlaValArgGluAspThrTyrProVal----- 428
      3468 -----CCAGAGAGAGCT-----TATTGAGTCCCTGCTCTCAGTGA 3503
QY      429 -----GlyThrGInGlyValProSerProAlaLeuAlaGInGlyGlyProGln 444
      3504 CAACCAAGAATGAGAAATTGGAGAAAACCAACAGACGTTGCCCAACAGAACCTTG- 3562
QY      445 GlySerTrpArgPheLeuGlnTrp----- 452
      3563 ACCTCTGTGAGCAGTGTGACGGTGCMAATGCTAGGATCCAAAGCCTTGAAAGCCACGCTAG 3622
QY      453 ---AsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArgHis 471
      3623 AGAAACTTCTTACCAAGAGAGTAGTAGCTGAAGCAGCTGCGCTGACCTGG----- 3673
QY      472 TyrAspPheArgGlnSerCysTrpValArgAlaIleSerGInGlyAspGlnLeuAlaPro 491
      3674 -----AGGTGAGCGTGCCTCTGCAG----- 3697
QY      492 CysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThr 511
      3698 -----ATGCTGAGGAGCTGCGGAGGCAAGCGCCGCCAGACT 3739
QY      512 AspSerAspGlnGlyThrProPheArgAlaArgAspGlnPro 526
      3740 CCAGAGCCAGACTGCACCTCAGCTGAGCCCAAGCGCATTTGACCG 3784
Db
```

RESULT 12

```
US-09-248-796A-1133
; Sequence 1133, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1133
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1133
```

```
Alignment Scores:
Pred. No.:      0.000707      Length:      1056
Score:          172.00      Matches:        49
Percent Similarity: 47.18%      Conservative: 43
Best Local Similarity: 25.13%      Mismatches:   67
Query Match:    5.77%      Indels:        36
DB:              4      Gaps:          7
```

US-10-071-838-2 (1-549) x US-09-248-796A-1133 (1-1056)

```
QY      145 IleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg----- 160
      28 ATGAAAGAGATTGAAATCGTACTTTCCTGATATATATACTTAAATTCATCAATCAAA 87
QY      161 -----AspArgTyrGlyThrLysGlnArgGlu----- 169
      88 GGTGTGTTTTCATCCTTAGAACTTTGCACAAAGAAAAGAACAGAAACATTATTGATC 147
QY      170 -----LeuLeuHisIleLeuLeuAlaTyrGluGluTyrAsnProGluValGlyTyrCys 187
      148 AAGTCATTAGAAGAGATATTGGTTCCTTGACACAGCATCAACCAATAAGATATTGT 207
QY      188 ArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPhe 207
      208 CAATCATTTGAATTTTGGCAGGCTTTTATTATTGTTTCATGCAAGAAAAGAAAGCATTT 267
QY      208 TrpAlaLeuValGlnLeuLeuAlaSerGluArg-----HisSer 220
      268 TGGATGCTTGT-----ATATTACAGAAAGAAATTATCCAAAGTTTCATCTGCMAAT 321
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAsp----- 237
      322 CTAGAAGGTGTTTACTGATCAAGGGTTTAAATGTTATGTGTGTTAAAGATATATACC 381
QY      238 GInGlnGlnHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLys 257
      382 CAATTATGCGAAGTTTAAAGGAAAATTTGCATGTGTGAACCTTTCTGAAGACAAG--- 438
QY      258 AspLeuCysGlyGlnCysSerProLeu-----GlyCysLeuIleArgIle 272
      439 ---ATTTATCAAGATTAACTCCAGTTTACATTAGTCACTTCATCTTGGTTTCATGTCA 495
QY      273 LeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlu 292
      496 TTGTGGGGAATTTACCTATTGAACAACTTTAAGATTATGAGATTACTTGTGATGAA 555
QY      293 GlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysVal 307
      556 GGTTCGAAAACAATTTTCAGATTCTCTTAAACCATATTCAAAATG 600
Db
```

RESULT 13

```
US-09-248-796A-1820
; Sequence 1820, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1820
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1820
```

```
Alignment Scores:
Pred. No.:      0.0284      Length:      1659
Score:          154.00      Matches:        46
Percent Similarity: 44.84%      Conservative: 54
Best Local Similarity: 20.63%      Mismatches:   111
Query Match:    5.16%      Indels:        12
DB:              4      Gaps:          4
```

US-10-071-838-2 (1-549) x US-09-248-796A-1820 (1-1659)

QY 78 TrpValAspMetLeuGlyAspTrpGluLysTyrLysSerSerArg-----LysLeuIle 95

Db 307 TGGAGTTGGTGTAAATGATTATGATGTTGTTAAACCATGAGCTGATTAATTAAAC 366
QY 96 AsparGAlaTyrIysGlyMetProMetAsnIleArgGlyProMetTrpSerValIleu 115
Db 367 CAGTGATATAATTCAGGAATACCAAGAGTTGAGGATTAATTTGGCACTAGTGCC 426
QY 116 AsnIleGluMetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLys 135
Db 427 AAATCCAAAGATTCCTCACTGAGAAGAT-----TTCTACCGCCCACTAAAA 471
QY 136 LysArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLys 155
Db 472 CTAGATCGTCATCCATCGAATAAGGAATCAAGAGACTTGACT-----AGAAC 522
QY 156 HisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeuHisIleLeu 175
Db 523 AGTTTCTTACCAATGTCGAGCGGTTAGCAAAAGCGATGAGCTTTCAATGTCATCAA 582
QY 176 AlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAla 195
Db 583 GCGTACTCTTGATGACCCCGATGGGCTACACCCAGGATGATTTTATTGACAGTG 642
QY 196 LeuPheLeuTyrLeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuAla 215
Db 643 CCATTGATATGATATGAACGAATCGAGTGTGTTTGTATTGTCACGTTAATG--- 699
QY 216 SerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 235
Db 700 ---AAAGATAACGGGCTCGCATTTGTTTGGCCGAAATGAAGGCTTCAGTGTG 756
QY 236 GlnAspGlnGlnHisValAlaIleHisSerGlnProLysThrMetGlyHisGlnAsp 255
Db 757 CTTTACGAATTGATCGATTGTTGAATCGTATTCTCAGTGTATACATCATTTTGGC 816
QY 256 LysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAsp 275
Db 817 AAACAAGTATCAAGTCGTCATGTCAGCTTCCAGTGTATACATCATTTTGGC 876
QY 276 GlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGln 295
Db 877 AAGTCCCTGACATGTTTAAAGAAATTACGACATTAATTGTACCCAGGGGATGAG 936
QY 296 AlaLeuMet 298
Db 937 TCCATATTG 945

RESULT 14
US-09-252-991A-71/c
; Sequence 71, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 71
; LENGTH: 7527
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-71

Alignment Scores:
Pred. No.: 0.547
Score: 149.00
Percent Similarity: 33.13%

Length: 7527
Matches: 121
Conservative: 46

Best Local Similarity: 24.01% Mismatches: 195
Query Match: 4.99% Indels: 142
DB: 4 Gaps: 22
US-10-071-838-2 (1-549) x US-09-252-991A-71 (1-7527)
QY 105 AsnIleArgGlyProMetTrpSerValLeuAsnIleGluMetLysLeuLysAsn 124
Db 5901 AATATCCAGACTGCCCTGATAGCGCTTCAACAGACACCTGGAACGAGCCAAAGAGCAT 5842
QY 125 -----ProGlyArgTyrGlnIleMetLysGlyLysArg 137
Db 5841 GAACGGGTGACACCTCTCCCGGGCGGAGCCAGCAGCTCTCCGATATGGCACC 5782
QY 138 SerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysIle 157
Db 5781 ---CAGCGCATGTCAGACGCTGCGCGCTCC---GCTGGCTGGGCGGCGCACCCG 5728
QY 158 Phe-----PheArgAspArgTyrGlyThrLysGlnArgGluLeuHisIleLeu 174
Db 5727 GTCCGGCGGCACTTCAGAGACCG-----CTG 5701
QY 175 LeuAlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIle--- 193
Db 5700 CTCGGCGCCAGCGCTACCGCCAGTAATCAAGTCCGCGCGCCCTCGCGCTGTCCAG 5641
QY 194 ---AlaIleLeuPheLeuTyrLeuProGluLysAspAlaPheTrpAlaLeuValGln 212
Db 5640 CCAGCGCGATGCCAGGCAATAGTGGCGTACTGACGCTCAATGCGCCAGCTCGG 5581
QY 213 LeuLeuAlaSerGluArgHisSerLeu-----GlnGlyPheHis 225
Db 5580 TTGTTGCG 5521
QY 226 SerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnHisValAlaIleHis 245
Db 5520 CGACCAACCGTCGACACGATATGCTGCTGGGTGATGACCAACACATGCTCTGCCAGC 5461
QY 246 SerGlnProLysThrMetGlyHisGlnAspLys-----LysAsp 258
Db 5460 CAGCGCAGCAGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5401
QY 259 LeuCysGlyGlnCysSerProLeuGlyCysLeuIleArg-----IleLeuIleAsp 275
Db 5400 TTCCTCGGCAACCGCTGGCGCAACGTTGCTTCCCTCGCGCGCGCGCGCGCGCGCG 5341
QY 276 Gly-----IleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuVal 291
Db 5340 AATGCGCAACGCGATGTTGCCAGATCGTCTGCGCGCGCGCGCGCGCGCGCGCG 5281
QY 292 GluGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArg 311
Db 5280 GCGGTGCGCAAGGTCTGCTGGCGCAGACCAACCAATGCAAGGCTGTCGCAACGCCG 5221
QY 312 Leu-ThrLysThrSerArg---CysGlyPro-TrpAlaArgPheCysAsnArgPheVal 330
Db 5220 CTGGTCCAGACGCGCACGACGTGACGTACGCTGGGAGATGATAGCGCGCTTCAGG 5161
QY 330 spThrTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLys 350
Db 5160 CTCAGTTTCCAGAGAACCAACATGCTGCTGAGCATGCGCAAAAGGCAAGCTCGGCAC 5101
QY 350 eutThr-----ArgLysGlnGlyAspLeuProProAlaLysProGluGlnLys 367
Db 5100 TCGCGCAATACTTGCAGCAGGCGGCAACGCTCGCGCGCGCGCGCGCGCGCGCGCG 5047
QY 367 erSerAlaSerArgProValProAlaSerArgGlyGlyThrLeuCysLysGlyAsp 387
Db 5046 AGCGGCGAAGTCTGCCAGACCGCGCGCTCGAACAAGG-----TGCGCAAGGCTAC 4996
QY 387 rgGlnAlaPro---ProGlyProProAlaArgPheProArg----- 399
Db 4995 GTCCAGCTCCAGATCTGCGCGCAGCGCGAGACCAACCGGCTGCGTGAAGGAGTGACC 4936

QY 400 -----ProIleTyrSerAlaSerProPro 408
Db 4935 GCCGAGTTCGAAGAAGTCTGTCGCGGGCGATGCCCTCGACACCGAGCAGCGCTCCCA 4876
QY 408 rGAlaProArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProV 428
Db 4875 GATCGCTGCCGACCGAAGCTTCGAGCCCT----- 4847
QY 428 aIGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpA 448
Db 4846 -----CCCGCGAGGCTCTCCGGCTGTGGCGGCGAG----- 4814
QY 448 rGpHeuGlnTrpAsnSerMetProAArgLeuProThrAspLeuAspValGluGlyProT 468
Db 4813 -----CCGCGTCCACCTTCG----- 4799
QY 468 rPpHeuArgHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspG 488
Db 4799 ----- 4799
QY 488 InLeuAlaProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAla-----P 506
Db 4798 --GCAGCGCCTTGC-----GATCCAGCTTGCCGTTGCGCTTCAGCGGCAAGCT 4753
QY 506 heAlaAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnP 526
Db 4752 GGACAGCACCCTGCCAAGCCGTGTCATGTAACGACGCAAGCGTCCGCCAGCAGGT 4693
QY 526 roCyAlaProThrSerGlyProCysLeuCysGlyLeuHisLeuGlnUserSerGlnPheP 546
Db 4692 GCGCAGCTCCGCGCAGAGGTCT-----CACCCCGCATAGCTCAGCGCC 4648
QY 546 roPProGly 548
Db 4647 CACCAAGT 4640

RESULT 15
US-09-919-039-243/c
; Sequence 243, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 243
; LENGTH: 3763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 346209.3
US-09-919-039-243

Alignment Scores:
Pred. No.: 0.227 Length: 3763
Score: 148.50 Matches: 95
Percent Similarity: 33.87% Conservative: 31
Best Local Similarity: 25.54% Mismatches: 147
Query Match: 4.98% Indels: 99
DB: 4 Gaps: 19

US-10-071-838-2 (1-549) x US-09-919-039-243 (1-3763)

QY 210 LeuValGlnLeuAlaSerGluArg--HisSerLeuGlnGlyPheHis--SerPro 227
Db 1133 CTCCTCAGCTGTCTCAGCAGCCCGGCTGCTCAGCTCCATGGCTCGACCCCGGAGACC 1074

QY 228 AsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnHisValAlaThrSerGln 247
Db 1073 CAAGGCCAATTTCGCCCTGTCTCTCTCCAGACAG----- 1038
QY 248 ProLysThrMetCysHisGlnAspLysAspLeuCysGlyGlnCysSerProLeuGly 267
Db 1037 -----CTCCTGGGCATCGTGCAGCTGGGCTTC 1011
QY 268 CysLeuIleArgIleLeu-----IleAspGlyIleSerLeuGlyLeuThrLeuArg-- 284
Db 1010 TGTGCTGTCACTCTCTGCTTAAGACGATGGTTTGAGACTCAGCTCGTTCAGCGCCCC 951
QY 285 -----LeuTrpAspValTyrLeuValGlnGlyGlnAlaLeu 297
Db 950 AGACACATTCTCCAGTTCAAGCTGGGCTGCTGCAGCTTCTAGCAGCCTCCGCTGTC 891
QY 298 ---MetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeu-ThrLysThrSe 316
Db 890 CCTCTCCCATCAACAGCCCGGCTGCACCTCTGCAGCTGTAACTCCAGCGCGCCT 831
QY 316 rArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspG 336
Db 830 CCGCTGTCAACCTCTCTGACGTGCAAGTTCAGAGGCTGCTCAGTTCT--GCCGCAAGCTC 774
QY 336 uAspThrValLeuLysHisLeu-ArgAlaSerMetLysLysLeuThrArgLysGlnGlyA 356
Db 773 GGACACCTCGGCTCCAGGCGCAGCGGCTTCTCCCATGACCTTTCGCCCTCCGGGC 714
QY 356 sPLeuPro-----ProProAlaLysProGlnGlnGly--SerSerAlaSerArgProV 373
Db 713 CTGCTCCAGCTGTCTCCGACAGTCCCGCAGGCGCTGGCGGCTGCTCAGCTCTCTG 654
QY 373 alProAlaSerArgGlyGlyLysThrLeuCysLysGly-----AspArgGlnAlaP 390
Db 653 CACTGCCGCTCGTGA-----TGCAGTCTCTCTCCAGAGTCTTCTTTCAG 606
QY 390 roPProGlyProProAlaArgPheProArgProIleTyrSerAlaSerPro----- 406
Db 605 CTCCTCAGCTCTCTGT-----CCCTCTTGACCGAGCTCTCTGTGCTG 558
QY 407 -----ProArgAlaProArgSerSerThrProCysP 417
Db 557 GGTGAGTCCAGCGGTCTCTCCAGCTGCGCCCGCAGCGCTCCAGCTCTCGCCAGGTC 498
QY 417 roGlyGlyAlaValArgGluAspThrTyrPro--ValGlyThrGlnGlyValProSerP 436
Db 497 CCGGCGCTGTCTCCGCTGTGCTGTCGACACAGCTCAGACTCCAGGCTCTCTGGGC 438
QY 436 ro-----AlaLeuAlaGlnGlyProGlnGlySerT 447
Db 437 CTCGGCCAGGCTGCTTGAGCCTCCCGCAGGATTTACAGCAGCTGGGCCCGCCACAC 378
QY 447 rPArgPheLeuGlnTrpAsnSerMetProAArgLeuProThrAspLeuAspValGluGlyP 467
Db 377 CTGCTCTTCTGCCCTGGCCAGGCGCAGCTCAGCTCTCTCTCCGCCAGCTGGGC 318
QY 467 roTrpPheArgHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluA 487
Db 317 CC-----GCAGCTCTCTGCCCCGCTGTGCTGCTCCAC 285
QY 487 sPGLnLeuAlaProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheA 507
Db 284 CATCTGCTCTCGAGCTGTGAGCTTCCCATCCAGC-----TCCGCTTTCAG 237
QY 507 laAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProC 527
Db 236 CTCTCCAGCTCTGAC-----GACCTTT 213
QY 527 ySAlaProThrSerGly-----ProCysLeu 535
Db 212 CTCCTCCTTCGTAAGCGGCTCTCTCAATGTCTG 181

Mon Feb 7 06:58:11 2005

us-10-071-838-2.rni

Page 16

Search completed: February 4, 2005, 11:53:57
Job time : 198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2005, 23:12:37 ; Search time 86 Seconds
(without alignments)
2290.028 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWMAQEREDILM.....TSGPCLCGHLHSSQFPFPGF 549

Scoring table: BLOSUM62

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2983	100.0	549	5	ABG70736	Abg70736 Human PRC
2	2983	100.0	549	7	ADC37383	Adc37383 Nuclear f
3	2978	99.8	549	3	AAy84901	AAY84901 A human p
4	2975	99.7	549	7	ADC37385	Adc37385 Nuclear f
5	2963	99.3	549	4	ABU53234	Abu53234 Human cel
6	2934.5	98.4	610	5	ABG70737	Abg70737 Human PRC
7	2850	95.5	531	5	ABP69593	Abp69593 Human pol
8	2835	95.0	527	7	ADC37389	Adc37389 Nuclear f
9	2827	94.8	527	5	ABP69592	Abp69592 Human pol
10	2827	94.8	527	7	ADC37387	Adc37387 Nuclear f
11	2827	94.8	527	7	ADM04434	Adm04434 Human pro
12	2784	93.3	549	5	ABG97507	Abg97507 Human NOV
13	2773	93.0	515	5	ABG70738	Abg70738 Human PRC
14	2191.5	73.5	499	4	ABU53235	Abu53235 Human cel
15	2191.5	73.5	805	4	ABG24026	Abg24026 Novel hum
16	2191.5	73.5	1388	4	ABG21344	Abg21344 Novel hum
17	2146.5	72.0	1066	4	ABG24025	Abg24025 Novel hum
18	2114	70.9	456	7	ADC37381	Adc37381 Nuclear f
19	2106	70.6	456	7	ADC37379	Adc37379 Nuclear f
20	1826	61.2	1142	4	ABG24027	Abg24027 Novel hum
21	1795.5	60.2	1131	4	ABG21347	Abg21347 Novel hum
22	1756	58.9	945	4	ABG19822	Abg19822 Novel hum
23	1756	58.9	945	4	ABG23335	Abg23335 Novel hum
24	1610	54.0	945	4	ABG22842	Abg22842 Novel hum
25	1550.5	52.0	805	4	ABG23336	Abg23336 Novel hum

26	1544.5	51.8	376	2	AAR90541	AAR90541 pTG4-5-CD
27	947	31.7	182	4	ABG23333	Abg23333 Novel hum
28	700	23.5	139	4	ABG15475	Abg15475 Novel hum
29	654.5	21.9	1127	4	AAM40477	Aam40477 Human pol
30	654.5	21.9	1127	4	AAM40478	Aam40478 Human pol
31	654.5	21.9	1127	7	ADC32650	Adc32650 Human nov
32	652.5	21.9	1125	4	ABG21348	Abg21348 Novel hum
33	652	21.9	1085	4	AAM38692	Aam38692 Human pol
34	652	21.9	1089	2	AAW82396	Aaw82396 Human UBP
35	652	21.9	1089	4	AAM38691	Aam38691 Human pol
36	652	21.9	1089	4	ABG23332	Abg23332 Novel hum
37	621.5	20.8	294	4	ABG15476	Abg15476 Novel hum
38	544	18.2	169	4	ABG24023	Abg24023 Novel hum
39	455	15.3	457	4	ABB67288	Abb67288 Drosophila
40	455	15.3	571	4	ABB63679	Abb63679 Drosophila
41	433.5	14.5	210	7	ADM05761	Adm05761 Human pro
42	425.5	14.3	185	4	ABG23334	Abg23334 Novel hum
43	387	13.0	80	4	ABG23331	Abg23331 Novel hum
44	337	11.3	336	4	ABG03482	Abg03482 Novel hum
45	323.5	10.8	356	3	AAG20059	Aag20059 Arabidops

ALIGNMENTS

RESULT 1	
ID	ABG70736 standard; protein; 549 AA.
XX	ABG70736;
AC	
XX	28-NOV-2002 (first entry)
DT	
XX	Human PRC17 protein.
DE	
XX	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;
KW	breast cancer; cytostatic; chromosome 17q11-12.
KW	
XX	Homo sapiens.
OS	
XX	WO200262958-A2.
PN	
XX	15-AUG-2002.
PD	
XX	
XX	08-FEB-2002; 2002WO-US003457.
PE	
XX	08-FEB-2001; 2001US-0267615P.
PR	
XX	(TULA-) TULARIK INC.
PA	
XX	Li J, Powers S, Xiang P, Peng Y;
PI	
XX	WPI; 2002-706902/76.
DR	N-PSDB; ABS54706.
DR	
XX	Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT	to detect diseases or conditions associated with altered PRC17 activity
PT	or expression relative to normal, for example cancer.
PT	
XX	Claim 2; Fig 5; 78bp; English.
PS	
XX	The present invention relates to a new PRC17 polypeptide. The invention
CC	is useful for detecting cancer cells (such as prostate tissue, breast
CC	tissue, lung tissue, ovarian tissue) in a biological sample. The
CC	invention is further useful for monitoring the efficacy of a therapeutic
CC	treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC	breast cancer). The molecules of the invention are useful diagnostically
CC	or prognostically to detect diseases or conditions associated with
CC	altered PRC17 activity or expression relative to normal, for example
CC	cancer. The present amino acid sequence represents the human PRC17
CC	protein of the invention. This sequence is encoded by the human PRC17
CC	gene located on chromosome 17q11-12.
CC	
XX	

Seq Sequence 549 AA;

Query Match 100.0%; Score 2983; DB 5; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.4e-257;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEGKGRAGLPEDKGPFRSYNNVNDHLGIVHETELPPL 60
DB 1 MDVVEVAGSWMWAQEREDIMKYEGKGRAGLPEDKGPFRSYNNVNDHLGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKWVMDLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSKWVMDLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLNKPGRYQIMKEKGRSSEHIQIRIDVSGTLRKHIFPRDRYGTQKORELHILLAYEY 180
DB 121 KLNKPGRYQIMKEKGRSSEHIQIRIDVSGTLRKHIFPRDRYGTQKORELHILLAYEY 180
QY 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGSTVQGLQDOQE 240
DB 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGSTVQGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSQPKTMGHQDKDLGGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVOQKRLTSTRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLPPL 360
DB 301 TRIAFKVOQKRLTSTRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLPPL 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRIWSASPPRAPRSSSTPCPGGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRIWSASPPRAPRSSSTPCPGGA 420
QY 421 VREDTYPVGTQGVSPALAAQGGPQGSWRFLOWNSMPRLPTDLDVEGPMFRHYDFRQSCWV 480
DB 421 VREDTYPVGTQGVSPALAAQGGPQGSWRFLOWNSMPRLPTDLDVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDQLAPCWQAEHPAERVRSAPAASTDSQGTFFRARDEQPCAPTSGPCLGHL 540
DB 481 RAISOEDQLAPCWQAEHPAERVRSAPAASTDSQGTFFRARDEQPCAPTSGPCLGHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 2

ADC37383
ID ADC37383 standard; protein; 549 AA.

AC ADC37383;

XX 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 216.

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
cancer; infectious disease; bone disease; AIDS;

KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
Neuroprotective; Nootropic; Cardiant; Gene therapy; human.

OS Homo sapiens.

XX WO2003048202-A2.

XX 12-JUN-2003.

XX 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
XX (ASAH) ASAH KASEI KK.

XX Matsuda A, Muramatsu S;

XX WPI, 2003-505282/47.

DR N-PSDB; ADC37382.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
useful for treating inflammation, autoimmune diseases, cancers,
infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
ischemic disorders.

PS Claim 1; SEQ ID NO 216; 938bp; English.

CC The present invention relates to novel proteins and their coding
sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
kappaB). The proteins and their coding sequences are useful for treating
a disease associated with NF-kappaB activation, such as inflammation,
autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
neurodegenerative diseases, or ischaemic disorders.

XX Sequence 549 AA;

Seq Query Match 100.0%; Score 2983; DB 7; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.4e-257;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEGKGRAGLPEDKGPFRSYNNVNDHLGIVHETELPPL 60
DB 1 MDVVEVAGSWMWAQEREDIMKYEGKGRAGLPEDKGPFRSYNNVNDHLGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKWVMDLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSKWVMDLGDWEKYSRKLIIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLNKPGRYQIMKEKGRSSEHIQIRIDVSGTLRKHIFPRDRYGTQKORELHILLAYEY 180
DB 121 KLNKPGRYQIMKEKGRSSEHIQIRIDVSGTLRKHIFPRDRYGTQKORELHILLAYEY 180
QY 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGSTVQGLQDOQE 240
DB 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGSTVQGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSQPKTMGHQDKDLGGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVOQKRLTSTRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLPPL 360
DB 301 TRIAFKVOQKRLTSTRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLPPL 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRIWSASPPRAPRSSSTPCPGGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRIWSASPPRAPRSSSTPCPGGA 420
QY 421 VREDTYPVGTQGVSPALAAQGGPQGSWRFLOWNSMPRLPTDLDVEGPMFRHYDFRQSCWV 480
DB 421 VREDTYPVGTQGVSPALAAQGGPQGSWRFLOWNSMPRLPTDLDVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDQLAPCWQAEHPAERVRSAPAASTDSQGTFFRARDEQPCAPTSGPCLGHL 540
DB 481 RAISOEDQLAPCWQAEHPAERVRSAPAASTDSQGTFFRARDEQPCAPTSGPCLGHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 3

AAV84901

ID	AA#	Standard	Protein	549 AA.
AC	XX	AA#84901		
DT	XX	21-AUG-2000	(first entry)	
DE	XX	A human proliferation and apoptosis related protein.		
KW	XX	Human; proliferation and apoptosis related protein; PROAP; psoriasis; cell proliferative disorder; immunological disorder; hepatitis; reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma; cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia; aschma; diabetes mellitus; osteoarthritis; endometriosis; uterine fibroid; menstrual cycle.		
XX	OS	Homo sapiens.		
FH	FT	Key	Location/Qualifiers	
FT	FT	Modified-site	61 /note= "potential phosphorylation site"	
FT	FT	Modified-site	73 /note= "potential phosphorylation site"	
FT	FT	Modified-site	90 /note= "potential phosphorylation site"	
FT	FT	Modified-site	91 /note= "potential phosphorylation site"	
FT	FT	Modified-site	98. .315 /note= "potential phosphorylation site"	
FT	FT	Domain	99 /note= "probable rabGAP domain"	
FT	FT	Modified-site	139 /note= "potential phosphorylation site"	
FT	FT	Modified-site	152 /note= "potential phosphorylation site"	
FT	FT	Modified-site	216 /note= "potential phosphorylation site"	
FT	FT	Modified-site	282 /note= "potential phosphorylation site"	
FT	FT	Modified-site	313 /note= "potential phosphorylation site"	
FT	FT	Modified-site	315 /note= "potential phosphorylation site"	
FT	FT	Modified-site	346 /note= "potential phosphorylation site"	
FT	FT	Modified-site	351 /note= "potential phosphorylation site"	
FT	FT	Modified-site	446 /note= "potential phosphorylation site"	
FT	FT	Modified-site	460 /note= "potential phosphorylation site"	
FT	FT	Modified-site	484 /note= "potential phosphorylation site"	
FT	FT	Modified-site	511 /note= "potential phosphorylation site"	
FT	FT	Modified-site	511 /note= "potential phosphorylation site"	
PN	XX	W0200023589-A2.		
PD	XX	27-APR-2000.		
PF	XX	19-OCT-1999;	99WO-US024511.	
PR	XX	20-OCT-1998;	98US-0172216P.	
PR	XX	04-FEB-1999;	99US-0118559P.	
PR	XX	11-FEB-1999;	99US-0172229P.	
PR	XX	22-APR-1999;	99US-0154336P.	
PA	XX	(INCY-) INCYTE PHARM INC.		
PI	XX	Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;		
PI	XX	Azmazai Y, Baughn MR, Yang J, Shih LL;		
DR	XX	WPI; 2000-339688/29.		
DR	XX	N-PSDB; AAA15001.		

Renewed
Experiment

Query	Match	Best Local Similarity	Score	DB	Length
Db	1 MDVVEVAGSWMAOEREDIMKEYEKGRAGLPEDKGPERSYNNVNDLGIYHETELPPL	99.8%	2978	DB 3	549
Qy	61 TAREAKQIRREISRSKSKVNDMLGDWEKYSSRKLIDRAYKGMNIRGPMWSVLNTEEM	99.8%	2978	DB 3	549
Db	61 TAREAKQIRREISRSKSKVNDMLGDWEKYSSRKLIDRAYKGMNIRGPMWSVLNTEEM	99.8%	2978	DB 3	549
Qy	121 KLNKPGRYQIMKEKGRSSSEHIQRIDRDVSGTLRKHIFFRDYGTQKRELHLILAYEY	99.8%	2978	DB 3	549
Db	121 KLNKPGRYQIMKEKGRSSSEHIQRIDRDVSGTLRKHIFFRDYGTQKRELHLILAYEY	99.8%	2978	DB 3	549
Qy	181 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLAERHSLQGFHSPNGTVOGLDQOE	99.8%	2978	DB 3	549
Db	181 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLAERHSLQGFHSPNGTVOGLDQOE	99.8%	2978	DB 3	549
Qy	241 HVAVATSQPKTMGHODKDLGQCSPLGLIRILIDGISLGLTLRLNDVYLVEGEQALMPI	99.8%	2978	DB 3	549
Db	241 HVAVATSQPKTMGHODKDLGQCSPLGLIRILIDGISLGLTLRLNDVYLVEGEQALMPI	99.8%	2978	DB 3	549
Qy	301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKLTKKQGLPPP	99.8%	2978	DB 3	549
Db	301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKLTKKQGLPPP	99.8%	2978	DB 3	549
Qy	361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARPPRIWSASPAPRPSSTPCPGA	99.8%	2978	DB 3	549
Db	361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARPPRIWSASPAPRPSSTPCPGA	99.8%	2978	DB 3	549
Qy	421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPELPTDLVEGPMFRHYDFRQSCWV	99.8%	2978	DB 3	549
Db	421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPELPTDLVEGPMFRHYDFRQSCWV	99.8%	2978	DB 3	549
Qy	481 RAISOEDQLAPCWAQEAHRAERVRSAFAAPSTDSDOGTFRARDEQPCAPTSGPCLGHL	99.8%	2978	DB 3	549
Db	481 RAISOEDQLAPCWAQEAHRAERVRSAFAAPSTDSDOGTFRARDEQPCAPTSGPCLGHL	99.8%	2978	DB 3	549
Qy	541 ESSQFPFPGF 549	99.8%	2978	DB 3	549
Db	541 ESSQFPFPGF 549	99.8%	2978	DB 3	549

AC ADC37385;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 218.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR N-PSDB; ADC37384.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 1; SEQ ID NO 218; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 549 AA;

Query Match 99.7%; Score 2975; DB 7; Length 549;
Best Local Similarity 99.8%; Pred. No. 7.2e-257;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPKEFRSYNNNVHDLGIHVHETELPPL 60
DB 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPKEFRSYNNNVHDLGIHVHETELPPL 60
QY 61 TAREAKQIRREISRKSKWVMDLGDWEKYKSRKLIIDRAYKGMPMNIRGPMWSVLNITEEM 120
DB 61 TAREAKQIRREISRKSKWVMDLGDWEKYKSRKLIIDRAYKGMPMNIRGPMWSVLNITEEM 120
QY 121 KLNKNGRYQIMKEKGKRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKORELHILAYEY 180
DB 121 KLNKNGRYQIMKEKGKRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKORELHILAYEY 180
QY 181 NPEVGYCRDLISHIAALFLYLPEEDAFWALVQLASERHSLQGFHSPNGTVOGLDQOE 240
DB 181 NPEVGYCRDLISHIAALFLYLPEEDAFWALVQLASERHSLQGFHSPNGTVOGLDQOE 240
QY 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDIGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDIGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVOQKRLTKTSRCGFWARFCNRPVDTWARDEDTVLKHLRASMKCLTRKQGDLP 360

DB 301 TRIAFKVOQKRLTKTSRCGFWARFCNRPVDTWARDEDTVLKHLRASMKCLTRKQGDLP 360
QY 361 AKPEQSSASRPVPASRGKTLCKDRQAPPGPARFPRIWSASPPRARSSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKDRQAPPGPARFPRIWSASPPRARSSTPCPGA 420
QY 421 VREDITYVGTQGVPSPALAQGGPQGSWRFLQNSMPRLPTDLDPVGPWFRHYDFRQSCWV 480
DB 421 VREDITYVGTQGVPSPALAQGGPQGSWRFLQNSMPRLPTDLDPVGPWFRHYDFRQSCWV 480
QY 481 RAISQEDQLAPCWOAEHPEAERVRSFAFAAPSTSDQGTFFRARDQPCAPTSGPCICGLHL 540
DB 481 RAISQEDQLAPCWOAEHPEAERVRSFAFAAPSTSDQGTFFRARDQPCAPTSGPCICGLHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 5
ABUS3234
ID ABUS3234 standard; protein; 549 AA.
XX
AC ABUS3234;
XX
DT 14-APR-2003 (first entry)
XX
DE Human cell cycle-associated protein from DKFzphhes3_35p22.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
DR N-PSDB; ABX71403.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Claim 21; Page 868; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a
CC polypeptide described in the disclosure of the invention
XX
SQ Sequence 549 AA;

Query Match 99.3%; Score 2963; DB 4; Length 549;
Best Local Similarity 99.3%; Pred. No. 8.5e-256;
Matches 545; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMWAQEREDIMKYEKGHRAGLPEDKGPKEFRSYNNNVHDLGIHVHETELPPL 60

Db 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKPFRRSYNNVNDHLGIVHETELPPL 60
Qy 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
Db 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNTEEM 120
Qy 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILLAYEEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILLAYEEY 180
Qy 181 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLQGFHSPNGGTWQGLQDOQE 240
Db 181 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLQGFHSPNGGTWQGLQDOQE 240
Qy 241 HVVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Qy 301 TRIAFKVQOKRLTITSRCGPMWARFCNRFVDTWARDDEDTVLKHLRASMKKLTRKQGLPP 360
Db 301 TRIAFKVQOKRLTITSRCGPMWARFCNRFVDTWARDDEDTVLKHLRASMKKLTRKQGLPP 360
Qy 361 AKPEQSSASRPVPASRGSKTLCKGDRQAPPGPARFPRPIWSASPPRAPRSSTPCPGGA 420
Db 361 AKPEQSSASRPVPASRGSKTLCKGDRQAPPGPARFPRPIWSASPPRAPRSSTPCPGGA 420
Qy 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 480
Qy 481 RAISOEDQLAPCWQAEHPAERVRSAFAAPSTDSQCTPRARDEQPCAPTSGPCLCGLHL 540
Db 481 RAISOEDQLAPCWQAEHPAERVRSAFAAPSTDSQCTPRARDEQPCAPTSGPCLCGLHL 540
Qy 541 ESSQFPFPGF 549
Db 541 ESSQFPFPGF 549

RESULT 6
ABG70737
ID ABG70737 standard; protein: 610 AA.
XX
AC ABG70737;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human PRC17 protein splice variant 1.
XX
KW Human; PRC17; prostate cancer; ovarian cancer; lung cancer;
XX breast cancer; cytostatic; chromosome 17q11-12; splice variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 141
FT /note= "Encoded by AT"
XX
PN WO200262958-A2.
XX
PD 15-AUG-2002.
XX
PF 08-FEB-2002; 2002WO-US003457.
XX
PR 08-FEB-2001; 2001US-0267615P.
XX
PA (TULA-) TULARIK INC.
XX
PI Li J, Powers S, Xiang P, Peng Y;
XX
DR WPI; 2002-706902/76.
DR N-PSDB; ABS54707.

XX Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.
XX
XX Claim 2; Page 64; 78pp; English.
PS
XX
CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present amino acid sequence represents the human PRC17
CC protein splice variant 1 of the invention. The human PRC17 gene is
CC located on chromosome 17q11-12
XX
SQ Sequence 610 AA;
Query Match 98.4%; Score 2934.5; DB 5; Length 610;
Best Local Similarity 89.8%; Pred. No. 3.5e-253;
Matches 548; Conservative 0; Mismatches 1; Indels 61; Gaps 1;
Qy 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKPFRRSYNNVNDHLGIVH----- 53
Db 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKPFRRSYNNVNDHLGIVOSCRSWES 60
Qy 54 -----ETELPP 59
Db 61 APQEGPCPFVPSPGLSPLEBRDRASPFGSAPRLGLQAPCSSSALPGLPYSETELPP 120
Qy 60 LTAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNIEE 119
Db 121 LTAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNIEE 180
Qy 120 MKLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILLAYEE 179
Db 181 MKLKNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHILLAYEE 240
Qy 180 YNPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLQGFHSPNGGTWQGLQDOQ 239
Db 241 YNPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLQGFHSPNGGTWQGLQDOQ 300
Qy 240 EHVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMP 299
Db 301 EHVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMP 360
Qy 300 ITRIAFKVQOKRLTITSRCGPMWARFCNRFVDTWARDDEDTVLKHLRASMKKLTRKQGLPP 359
Db 361 ITRIAFKVQOKRLTITSRCGPMWARFCNRFVDTWARDDEDTVLKHLRASMKKLTRKQGLPP 420
Qy 360 PAKPEQSSASRPVPASRGSKTLCKGDRQAPPGPARFPRPIWSASPPRAPRSSTPCPGG 419
Db 421 PAKPEQSSASRPVPASRGSKTLCKGDRQAPPGPARFPRPIWSASPPRAPRSSTPCPGG 480
Qy 420 AVREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 479
Db 481 AVREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCW 540
Qy 480 VRAISOEDQLAPCWQAEHPAERVRSAFAAPSTDSQCTPRARDEQPCAPTSGPCLCGLH 539
Db 541 VRAISOEDQLAPCWQAEHPAERVRSAFAAPSTDSQCTPRARDEQPCAPTSGPCLCGLH 600
Qy 540 LESSQFPFPGF 549
Db 601 LESSQFPFPGF 610

RESULT 7
ABP69593
ID ABP69593 standard; protein: 531 AA.

XX AC ABP69593;
XX DT 20-JAN-2003 (first entry)
XX DE Human polypeptide SEQ ID NO 1640.
XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic.
XX OS Homo sapiens.
XX PN WO200270539-A2.
XX PD 12-SEP-2002.
XX PF 05-MAR-2002; 2002WO-US005095.
XX PR 05-MAR-2001; 2001US-00799451.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2002-759812/82.
DR N-PSDB; ABZ11810.
XX
XX PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1640; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic), liver
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 531 AA;

Query Match 95.5%; Score 2850; DB 5; Length 531;
Best Local Similarity 96.2%; Pred. No. 1e-245;
Matches 528; Conservative 1; Mismatches 2; Indels 18; Gaps 1;

QY 1 MDVVEVAGSWMAQEREDIIIMKEKGRAGLPEDKGPKPFRSYNNNVNDHLGIWHETELPPL 60
DB 1 MDVVEVAGSWMAQEREDIIIMKEKGRAGLPEDKGPKPFRSYNNNVNDHLGIWHETELPPL 60
QY 61 TAREAKQIRREISRKSKWMDMLGDMWEKYSRKRLIDRAYKGMPMNIRGPMWSVLNTEEM 120
DB 61 TAREAKQIRREISRKSKWMDMLGDMWEKYSRKRLIDRAYKGMPMNIRGPMWSVLNTEEM 120
PT

QY 121 KLKNPGRYQIMEKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQRELLHLLAYEEY 180
DB 121 KL-----NIQRIDRVSGTLRKHIFFRDRYGTQRELLHLLAYEEY 162
QY 181 NPEVGYCRDLSHIAALFLVYLPEEDAFWALVOLLAASERHSLQGFSBNGGTVOGLDQOE 240
DB 163 NPEVGYCRDLSHIAALFLVYLPEEDAFWALVOLLAASERHSLQGFSBNGGTVOGLDQOE 222
QY 241 HVVATSOPTMGHQDKDLCGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 300
DB 223 HVVATSOPTMGHQDKDLCGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPI 282
QY 301 TRIAFKYQQKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKLTKXGDLPPP 360
DB 283 TRIAFKYQQKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKLTKXGDLPPP 342
QY 361 AKPEQSSASRPVPASRGCKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 420
DB 343 AKPEQSSASRPVPASRGCKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 402
QY 421 VREDTYPVGTQGVPSPALAOGGPGSWRFLQWNSMPRLPTDLVEGPFWRHYDFRQSCWV 480
DB 403 VREDTYPVGTQGVPSPALAOGGPGSWRFLQWNSMPRLPTDLVEGPFWRHYDFRQSCWV 462
QY 481 RAISOEDQAPCWAHEHPAERVSAFAAPSTDSOGTPEFRARDEQPCAPTSGPCLCGHL 540
DB 463 RAISOEDQAPCWAHEHPAERVSAFAAPSTDSOGTPEFRARDEQPCAPTSGPCLCGHL 522
QY 541 ESSQFPFPGF 549
DB 523 ESSQFPFPGF 531

RESULT 8
ADC37389
ID ADC37389 standard; protein; 527 AA.
XX
AC ADC37389;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 222.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR N-PSDB; ADC37388.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.

XX Claim 1, SEQ ID NO 222, 938pp; English.
PS
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 527 AA;

Query Match 95.0%; Score 2835; DB 7; Length 527;
Best Local Similarity 95.8%; Pred. No. 2.3e-244;
Matches 526; Conservative 0; Mismatches 1; Indels 22; Gaps 1;

Qy 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKEFRSYNNNVNHLGIVHETELPPL 60
Db 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKEFRSYNNNVNHLGIVHETELPPL 60
Qy 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLDIDRAYKGMNIRGPMMSVLLNIEEM 120
Db 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLDIDRAYKGMNIRGPMMSVLLNIEEM 120
Qy 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIFPRDRYGTQKRELIHILAYEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIFPRDRYGTQKRELIHILAYEY 180
Qy 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFSHPNGGTVOGLQDOQE 240
Db 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFSHPNGGTVOGLQDOQE 240
Qy 241 HVAVTSQPKTMGHQDKDLGQCSPLGCLIRILIDIGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVAVTSQPKTMGHQ-----ISLGLTLRLMDVYLVEGEQALMPI 278
Qy 301 TRIAFKVQOKRLTTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 360
Db 279 TRIAFKVQOKRLTTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 338
Qy 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 420
Db 339 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRARSSTPCPGA 398
Qy 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVBGMPFRHYDFRQSCW 480
Db 399 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVBGMPFRHYDFRQSCW 458
Qy 481 RAISOEDQLAPCWAQAEHRAERVRSAFAAPSTDSQCTPFARDEQPCAPTSGPCLGTL 540
Db 459 RAISOEDQLAPCWAQAEHRAERVRSAFAAPSTDSQCTPFARDEQPCAPTSGPCLGTL 518
Qy 541 ESSQFPFPGF 549
Db 519 ESSQFPFPGF 527

RESULT 9
ABP69592
ID ABP69592 standard; protein; 527 AA.
XX
AC ABP69592;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1639.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiParkinsonian; antidiabetic; immunosuppressive; dermatological;

KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.
XX
XX Homo sapiens.
XX OS
XX WO200270539-A2.
XX
XX
XX 12-SEP-2002.
XX
XX 05-MAR-2002; 2002WO-US005095.
XX
XX 05-MAR-2001; 2001US-00799451.
XX
XX
XX (HYSE-) HYSEQ INC.
XX
XX
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2002-759812/82.
XX
XX N-PSDB; ABZ11809.

Claim 9; SEQ ID NO 1639; 1012pp + Sequence listing; English.

The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 527 AA;

Query Match 94.8%; Score 2827; DB 5; Length 527;
Best Local Similarity 95.6%; Pred. No. 1.2e-243;
Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;

Qy 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKEFRSYNNNVNHLGIVHETELPPL 60
Db 1 MDVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPKEFRSYNNNVNHLGIVHETELPPL 60
Qy 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLDIDRAYKGMNIRGPMMSVLLNIEEM 120
Db 61 TAREAKQIRREISRSKSKVNDMLGDWEKYKSSRKLDIDRAYKGMNIRGPMMSVLLNIEEM 120
Qy 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIFPRDRYGTQKRELIHILAYEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIFPRDRYGTQKRELIHILAYEY 180
Qy 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFSHPNGGTVOGLQDOQE 240
Db 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFSHPNGGTVOGLQDOQE 240
Qy 241 HVAVTSQPKTMGHQDKDLGQCSPLGCLIRILIDIGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVAVTSQPKTMGHQ-----ISLGLTLRLMDVYLVEGEQALMPI 278
Qy 301 TRIAFKVQOKRLTTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 360

Db 279 TRIAFKVOQKRLTKTSRCGPMWAFPCNRFVDIWARDEDIVLKHILRASMKCLTRKQGDLP 338
QY 361 AKPEOGSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGGA 420
Db 339 AKPEOGSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGGA 398
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVDEGPMFRHYDFRQSCWV 480
Db 399 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVDEGPMFRHYDFRQSCWV 458
QY 481 RAISOEDQLAPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 540
Db 459 RAISOEDQLAPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 518
QY 541 ESSQFPFPGF 549
Db 519 ESSQFPFPGF 527

RESULT 10
ADC37387
ID ADC37387 standard; protein; 527 AA.
AC ADC37387;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 220.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Antinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR N-PSDB; ADC37386.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 1; SEQ ID NO 220; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
SQ Sequence 527 AA;

Query Match 94.8%; Score 2827; DB 7; length 527;

Best Local Similarity 95.6%; Pred. No. 1.2e-243;
Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
QY 1 MDVVEAGSWWAQEREDIMKEYEKGRAGLPEDKPKPFRSYNNVNDHLGIVHETELPPL 60
Db 1 MDVVEAGSWWAQEREDIMKEYEKGRAGLPEDKPKPFRSYNNVNDHLGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKRWMDLGDWEKYSRKLIDRAYKGMNINRGPMWSVLTNTEEM 120
Db 61 TAREAKQIRREISRSKRWMDLGDWEKYSRKLIDRAYKGMNINRGPMWSVLTNTEEM 120
QY 121 KLNKPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELHILAYEY 180
Db 121 KLNKPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELHILAYEY 180
QY 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOE 240
Db 181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGGTVOGLQDOE 240
QY 241 HVVATSOPTMGHODKDLGCGCSPLGCLIRILIDIGISGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSOPTMGHQ-----ISGLTLRLMDVYLVEGEQALMPI 278
QY 301 TRIAFKVOQKRLTKTSRCGPMWAFPCNRFVDIWARDEDIVLKHILRASMKCLTRKQGDLP 360
Db 279 TRIAFKVOQKRLTKTSRCGPMWAFPCNRFVDIWARDEDIVLKHILRASMKCLTRKQGDLP 338
QY 361 AKPEOGSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGGA 420
Db 339 AKPEOGSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGGA 398
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVDEGPMFRHYDFRQSCWV 480
Db 399 VREDTYPVGTQGVSPALAQGGPQGSWRFLOWNSMPRLPTDLVDEGPMFRHYDFRQSCWV 458
QY 481 RAISOEDQLAPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 540
Db 459 RAISOEDQLAPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 518
QY 541 ESSQFPFPGF 549
Db 519 ESSQFPFPGF 527

RESULT 11
ADM04434
ID ADM04434 standard; protein; 527 AA.
XX
AC ADM04434;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein of the invention SEQ ID NO:3119.
XX
KW human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
DR WPI; 2003-723558/69.

DR N-PSDB; ADM01991.
XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 3119; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotide ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.
XX
SQ Sequence 527 AA;

Query Match 94.8%; Score 2827; DB 7; Length 527;
Best Local Similarity 95.6%; Pred. No. 1.2e-243;
Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;

QY 1 MDVVEAGSWWAQEREDIMKYEKGRAGLPEDKGPFRSYNNNVNHLGIHETELPPL 60
Db 1 MDVVEAGSWWAQEREDIMKYEKGRAGLPEDKGPFRSYNNNVNHLGIHETELPPL 60
QY 61 TAREAKQIRREISRSKSWNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
Db 61 TAREAKQIRREISRSKSWNDMLGDWEKYKSSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
QY 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIIPFRDRYGTQRELLHILAYEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQIRIDRVSGTLRKHIIPFRDRYGTQRELLHILAYEY 180
QY 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFHSPNGGTVOGLDQOE 240
Db 181 NPEVGYCRDLSHIALFLLYLPBEDAFWALVQLASERHSLQGFHSPNGGTVOGLDQOE 240
QY 241 HVVATSQPKTMGHODKKDLGQCSPLGCLIRILIDIGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHQ-----ISLGLTLRLMDVYLVEGEQALMPI 278
QY 301 TRIAFKVOQKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTRKQGDLP 360
Db 279 TRIAFKVOQKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTRKQGDLP 338
QY 361 AKPEQSSASRPVAPSRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGA 420
Db 339 AKPEQSSASRPVAPSRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGA 398
QY 421 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVVEGFWFRHYDFRQSCWV 480
Db 399 VREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVVEGFWFRHYDFRQSCWV 458
QY 481 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSQCTPFARARDEQCAPTSGPCLGHL 540
Db 459 RAISOEDQLAPCWAHEHPAERVRSAFAAPSTDSQCTPFARARDEQCAPTSGPCLGHL 518
QY 541 ESSQPPPGF 549
Db 519 ESSQPPPGF 527

RESULT 12
ABG97507
ID ABG97507 standard; protein, 549 AA.
AC ABG97507;
XX
DT 16-DEC-2002 (first entry)

XX DE Human NOVX26 protein.
XX
KW Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
KW Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
KW asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
KW bacterial infection; parasitic infection; graft-versus-host disease;
KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW angiogenesis.
XX
OS Homo sapiens.
XX
PN WO200272770-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US007283.
XX
PR 08-MAR-2001; 2001US-0274281P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-MAY-2001; 2001US-0288148P.
PR 31-MAY-2001; 2001US-0294821P.
PR 31-OCT-2001; 2001US-0335302P.
PR 04-DEC-2001; 2001US-0338375P.
PR 07-MAR-2002; 2002US-00094466.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Vernet CA, Tchernev VT, Malyankar UM, Gerlach VL;
PI Li L, Zernhusen BD, Paturajan M, Gusev VY, Kekuda R, Pena CEA;
PI Zhong M, Gangolli EA, Taupier RJ;
XX
XX WPI; 2002-713508/77.
DR N-PSDB; ABS78751.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
PT Parkinson's disease.
XX
XX Claim 1; Page 164; 266pp; English.
XX
XX The present invention relates to a new polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
CC polypeptides and antibodies are useful for treating, preventing or
CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,

CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC infections, or graft-versus-host disease. The nucleic acids and
CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVA substances for use in therapeutic or
CC diagnostic methods. The nucleic acids are further used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine, and
CC pharmacogenomics. The present amino acid sequence represents a human NOVX
XX protein of the invention

Sequence 549 AA;

Query Match 93.3%; Score 2784; DB 5; Length 549;
Best Local Similarity 93.8%; Pred. No. 8.7e-240;
Matches 515; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

OY 1 MDVEVAGSWWAQEREDIMKYEGHRAGLPEDKGPKEFRSYNNNVHGLIVHETELPPL 60
Db 1 MDVEVAGSWWAQEREDIMKYEGHRAGLPEDKGPKEFRSYNNNVHGLIVHETELPPL 60
OY 61 TAREAKQIRREISRSKSWVDMLGDWEKYKSRKIDRAYKGMPMNIRGPMWSVLNIEEM 120
Db 61 TAREVQKIRREISRSKSWVMWLGEMDYYKNSRKIDRAYQGIPIWIRGPMWSVLNIEEI 120
OY 121 KLNKPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIFFRDRYGTQKRELIHILAYEY 180
Db 121 KLNKPGRYQIMKEKGKRSSEHIQOMDLDSVGTLRHIFFRDRYGTQKRELIHILAYEY 180
OY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLDQOE 240
Db 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLDQOE 240
OY 241 HVAATSQPTMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVAATSLPNTMMHODKDLGQCSISGLILIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
OY 301 TRIAFKVOQKRLTKTSRCGPMWAFRCNRFVDTWARDEDVTLKHLRASMKLTKRQGDLPPL 360
Db 301 TRIAFKVOLERLTKTSRCGPMWAFRCNRFVDTWARDEDVTLKHLRASMKLTKRQGDLPPL 360
OY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSSTPCPGGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSSTPCPGGA 420
OY 421 VREDTYPVGTQGVSPALAQGGPQSWRFLQWNSMPRLPTDLVDEGWFPHYDFRQSCWV 480
Db 421 VREDTYPVGTQGVSPALAQGGPQSWRFLQWNSMPRLPTDLVDEGWFPHYDFRQSCWV 480
OY 481 RAISOEQOLAPCWAHEHPAERVRSAFAAPSTDSOGTTPRARDEQPCAPTSGPCLCGLHL 540
Db 481 RAISOEQOLAPCWAHEHPAERVRSAFAAPSTDSOGTTPRARDEQPCAPTSGPCLCGLHL 540
OY 541 ESSQPPPPGF 549
Db 541 ESSQPPPPGF 549

RESULT 13
ABG70738
ID ABG70738 standard; protein; 515 AA.
XX
AC ABG70738;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human PRC17 protein splice variant 2.
XX
KW Human; PRC17; prostate cancer; ovarian cancer; lung cancer;

KW breast cancer; cytostatic; chromosome 17q11-12; splice variant.
XX Homo sapiens.
OS
XX WO200262958-A2.
PN
XX 15-AUG-2002.
PD
XX 08-FEB-2002; 2002WO-US003457.
PF
XX 08-FEB-2001; 2001US-0267615P.
PR
XX (TULA-) TULARIK INC.
PA
XX
XX
PI Li J, Powers S, Xiang P, Peng Y;
PI
XX
DR WPI; 2002-706902/76.
DR N-PSDB; ABSS4708.
XX

Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.
XX
PS Claim 2; Page 65; 78bp; English.

The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present amino acid sequence represents the human PRC17
CC protein splice variant 2 of the invention. The human PRC17 gene is
CC located on chromosome 17q11-12

Sequence 515 AA;

Query Match 93.0%; Score 2773; DB 5; Length 515;
Best Local Similarity 93.8%; Pred. No. 7.7e-239;
Matches 515; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

OY 1 MDVEVAGSWWAQEREDIMKYEGHRAGLPEDKGPKEFRSYNNNVHGLIVHETELPPL 60
Db 1 MDVEVAGSWWAQEREDIMKYEGHRAGLPEDKGPKEFRSYNNNVHGLIVHETELPPL 60
OY 61 TAREAKQIRREISRSKSWVDMLGDWEKYKSRKIDRAYKGMPMNIRGPMWSVLNIEEM 120
Db 61 TAREAKQIRREISRSKSWVDMLGDWEKYKSRKIDRAYKGMPMNIRGPMWSVLNIEEM 120
OY 121 KLNKPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIFFRDRYGTQKRELIHILAYEY 180
Db 121 KLNKPGRYQIMKEKGKRSSEHIQIRIDRVSGTLRKHIFFRDRYGTQKRELIHILAYEY 180
OY 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLDQOE 240
Db 181 NPEVGYCRDLISHIALFLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLDQOE 240
OY 241 HVAATSQPTMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVAATSQPTMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
OY 301 TRIAFKVOQKRLTKTSRCGPMWAFRCNRFVDTWARDEDVTLKHLRASMKLTKRQGDLPPL 360
Db 301 TRIAFKVOQKRLTKTSRCGPMWAFRCNRFVDTWARDEDVTLKHLRASMKLTKRQGDLPPL 360
OY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSSTPCPGGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSSTPCPGGA 420
OY 421 VREDTYPVGTQGVSPALAQGGPQSWRFLQWNSMPRLPTDLVDEGWFPHYDFRQSCWV 480
Db 421 VREDTYPVGTQGVSPALAQGGPQSWRFLQWNSMPRLPTDLVDEGWFPHYDFRQSCWV 480

Db 387 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDL DVEGPFWRHYDFRQSCWV 446
QY 481 RAISQEDQLAPCWOAEHPAERVRSAFAAPSTDSDGTPFERARDEQPCAPTSGPCLGLHL 540
Db 447 RAISQEDQLAPCWOAEHPAERVRSAFAAPSTDSDGTPFERARDEQPCAPTSGPCLGLHL 506
QY 541 ESSQFPFPGF 549
Db 507 ESSQFPFPGF 515

RESULT 14
ABUS3235
ID ABUS3235 standard; protein; 499 AA.
AC ABUS3235;
XX
DT 14-APR-2003 (first entry)
DE Human cell cycle-associated DKFZphr3522 homologue.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Example III; Page 868-869; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
CC of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 499 AA;

Query Match 73.5%; Score 2191.5; DB 4; Length 499;
Best Local Similarity 81.6%; Pred. No. 7.6e-187;
Matches 408; Conservative 33; Mismatches 58; Indels 1; Gaps 1;

QY 1 MDVVEVAGSWMAQERDIIIMYEKGHRAGLPEDKGPFRSYNNNVHGIIVHETELPPL 60
Db 1 MDVVENADSLQAQERKDIIMKYDKGHRAGLPEDKGPBPV-GINSSIRFGILHETELPVP 59
QY 61 TAREAKQIRREISRSKRVNDMLGDWEKYKSKRLIDRAYKGMPPNIRGPMWSVLNIEEM 120
Db 60 TAREAKIRREMTIRISKWMEMLGEWETKHSKLI DRVYKGI PMNIRGPMWSVLNIQEI 119
QY 121 KLNKPGRYQIMKEKGKSSSEHIQIRIDRVSGTLRKHIIFRDRYGTQKQELLHILAYEEY 180
Db 120 KLNKPGRYQIMKERGKSSSEHIHIDLVRTTLRNHVFFRDRYGAQKRELFIYLAYSEY 179

QY 181 NPEVGYCRDLSHTAALFLYLPEBDAFWALVOLLASERHSLQGFHSPNGCTVQGLQDQE 240
Db 180 NPEVGYCRDLSHTALFLYLPEBDAFWALVOLLASERHSLPGFHSNGCTVQGLQDQE 239
QY 241 HVVATSQPKTMGHQDKDLQGCSPGLCLIRILIDIGISLGLTRLMVYLVGEQALMPI 300
Db 240 HVPFKSQPKTMHQDKEGLCGQCASTGLRLNIDIGISLGLTRLMVYLVGEQALMPI 299
QY 301 TRIAFKVOQKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKKLTRKQGLPPP 360
Db 300 TSIALKVQOKRLMKTSRCGLMARLRNQFPDTWAMNDTVLKHILRASTKKLTRKQGLPPP 359
QY 361 AKREGSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPRSSTPCPGA 420
Db 360 AKREGSLAPRPVPASRGKTLCKGYRQAPPGPPAQFORPICASBPWASRFSTPCPGA 419
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDL DVEGPFWRHYDFRQSCWV 480
Db 420 VREDTYPVGTQGVSPALAOGGPQGSWRFLEWKSMPRLPTDL DVGPPFPHYDFERSCWV 479
QY 481 RAISQEDQLAPCWOAEHPAE 500
Db 480 RAISQEDQLATCWOAEHCGE 499

RESULT 15
ABG24026
ID ABG24026 standard; protein; 805 AA.
XX
AC ABG24026;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24017.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88213.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 54385; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food

supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 805 AA;
SQ

Query Match	73.5%;	Score 2191.5;	DB 4;	Length 805;
Best Local Similarity	81.6%;	Pred. No. 1.5e-186;		
Matches 408; Conservative	33;	Mismatches 58;	Indels 1;	Gaps 1;

QY	1	MDVEVAGS	WMAOERED	IIMKYEK	GHAGL	PEDGPKP	FRSYNN	NVNDH	GIVHETEL	PPL	60		
Db	20	MDWVENAD	SLQAQER	KDILMKYD	KHRAGL	PEDKGPE	PV - GINS	SIDRE	GILHETEL	PPV	78		
QY	61	TAREAKOIR	REISRSK	SKVMDML	GDWEKYSS	RKLIDRAYK	GMPNIR	GEMSVL	LNIBEM	120			
Db	79	TAREAKKIR	REMTRTS	KMMEM	LGEWET	YKHSKLI	DRVYKGI	PMNIR	GPAWSV	LNIOEI	138		
QY	121	KLKNPGR	YOIMKEK	KRSSEHI	QRI	IDRVS	GTLRKHI	PFRDR	YGTQORE	LHILAYE	180		
Db	139	KLKNPGR	YOIMKER	GKRSSEH	IHHID	LVTR	LRNHV	FFRDR	YGAKORE	LYILLAYSE	198		
QY	181	NPEVGYCR	DLSHIAAL	FLLYL	PEEDA	FWALVOL	LASERHS	LOGFHS	PNGTVO	GLQDOOE	240		
Db	199	NPEVGYCR	DLSHITAL	FLLYL	PEEDA	FWALVOL	LASERHS	LPGFHS	PNGTVO	GLQDOOE	258		
QY	241	HVVATSGP	KTMGHQD	KDLCGQCS	PLGCLIR	ILIDG	ISLGLT	RLMDVYL	VEGEQAL	MP	300		
Db	259	HVVATSGP	KTMGHQD	KEGLCGQCS	ASLGLLE	NLIDG	ISLGLT	RLMDVYL	VEGEQV	LMPI	318		
QY	301	TRIAFKYQ	OKRLTKTS	RRCGP	WARFCNR	FVD	TWARD	EDTVL	KHLRAS	MKKLTKRQGD	PPP	360	
Db	319	TSIALKYQ	OKRLMKTS	RRCGL	MARLRN	QF	FTWAMND	DTV	LKHLRA	STKCLTKRQGD	PPP	378	
QY	361	AKPEQGS	SASRPVP	PASRG	KTLCKGDR	QAPG	PPARF	PRPI	WSAS	PPRARS	STPCPGGA	420	
Db	379	AKREGG	SLAPRPVP	PASRG	KTLCKGYR	QAPG	PPAQ	QRP	ICSAS	PPWASR	STPCPGGA	438	
QY	421	VREDTYP	VGTQGVPS	PALAAQ	GGPQGS	WRFL	OWNSM	PR	PTDLD	VEG	PWFRHYD	FRQSCWV	480
Db	439	VREDTYP	VGTQGVPS	LALAAQ	GGPQGS	WRFL	ENKSM	PR	PTDLD	IG	GPWFPHYD	FRSCWV	498
QY	481	RAISOED	QOLAPC	WQAEH	PAE	500							
Db	499	RAISOED	QOLATC	WQAEH	CGE	518							

Search completed: February 4, 2005, 05:08:18
Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 4, 2005, 05:10:13 ; Search time 84 Seconds
(without alignments)
2361.284 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWMAQEREDIIIM.....TSGPCLGLHLSQFPFPGF 549

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2983	100.0	549	US-10-071-838-2	Sequence 2, Appli
2	2978	99.8	549	US-10-839-882-12	Sequence 12, Appl
3	2952	99.0	549	US-10-071-838-18	Sequence 18, Appl
4	2934.5	98.4	610	US-10-071-838-4	Sequence 4, Appli
5	2827	94.8	527	US-10-108-260A-3119	Sequence 3119, Ap
6	2784	93.3	549	US-10-094-466-62	Sequence 62, Appl
7	2773	93.0	515	US-10-071-838-6	Sequence 6, Appli
8	2191.5	73.5	786	US-10-071-838-17	Sequence 17, Appl
9	433.5	14.5	210	US-10-108-260A-4446	Sequence 4446, Ap
10	348.5	11.7	381	US-10-425-115-237632	Sequence 237632,
11	348.5	11.7	434	US-10-425-114-64834	Sequence 64834, A
12	348.5	11.7	434	US-10-425-114-70559	Sequence 70559, A
13	329	11.0	433	US-10-437-963-139162	Sequence 139162,

14	321.5	10.8	537	15	US-10-051-874-138	Sequence 138, App
15	320.5	10.7	533	15	US-10-051-874-137	Sequence 137, App
16	320.5	10.7	533	16	US-10-408-765A-1052	Sequence 1052, Ap
17	305.5	10.2	500	15	US-10-051-874-135	Sequence 135, App
18	299.5	10.0	508	15	US-10-051-874-46	Sequence 46, Appl
19	299.5	10.0	508	15	US-10-051-874-134	Sequence 134, App
20	298	10.0	438	15	US-10-051-874-136	Sequence 136, App
21	290.5	9.7	241	16	US-10-767-701-32143	Sequence 32143, A
22	282	9.5	289	15	US-10-424-599-148597	Sequence 148597,
23	266.5	8.9	209	15	US-10-051-874-139	Sequence 139, App
24	231	7.7	44	9	US-09-864-761-43564	Sequence 43564, A
25	229.5	7.7	811	16	US-10-437-963-118949	Sequence 118949,
26	226.5	7.6	723	16	US-10-322-281-455	Sequence 455, App
27	226	7.6	806	16	US-10-322-281-458	Sequence 458, App
28	222	7.4	607	17	US-10-370-715B-328	Sequence 328, App
29	214	7.2	468	14	US-10-258-883-2	Sequence 2, Appli
30	214	7.2	924	14	US-10-177-293-244	Sequence 244, App
31	214	7.2	928	11	US-09-989-890-198	Sequence 198, App
32	214	7.2	928	16	US-10-473-576-16	Sequence 16, Appl
33	214	7.2	1004	14	US-10-094-466-58	Sequence 58, Appl
34	214	7.2	1026	16	US-10-467-595-5	Sequence 5, Appli
35	209	7.0	768	9	US-09-764-868-728	Sequence 728, App
36	209	7.0	768	11	US-09-764-875-922	Sequence 922, App
37	202.5	6.8	815	15	US-10-108-260A-3374	Sequence 3374, Ap
38	201.5	6.8	775	15	US-10-112-944-291	Sequence 291, App
39	201.5	6.8	775	15	US-10-112-944-749	Sequence 749, App
40	200	6.7	917	14	US-10-221-658-4	Sequence 4, Appli
41	199.5	6.7	814	15	US-10-158-057-243	Sequence 243, App
42	198	6.6	756	11	US-09-764-875-847	Sequence 847, App
43	197.5	6.6	366	15	US-10-264-237-1756	Sequence 1756, Ap
44	197.5	6.6	373	9	US-09-764-848-32	Sequence 32, Appl
45	197.5	6.6	373	14	US-10-116-016-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-10-071-838-2
; Sequence 2, Application US/10071838
; Publication No. US2003004481A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17
US-10-071-838-2

Query Match 100.0%; Score 2983; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 6.4e-225;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MDVVEVAGSWMAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVVDHIGIVHETELPPL 60
Db 1 MDVVEVAGSWMAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVVDHIGIVHETELPPL 60
Cy 61 TAREAKQIRRETSRKSXWVDMIGDNEKYKSRKLIIDRAYKGMNINRGPMWSVLNTEEM 120

Db 61 TAREAKQIRREISRSKSKWVMDLGDWEKYKSRKLIIDRAYKGMNIRGPMWSVLNTEEM 120
QY 121 KLNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFPRDRYGTQORELHILLAYEY 180
Db 121 KLNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFPRDRYGTQORELHILLAYEY 180
QY 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOQE 240
Db 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGQCCSPGLCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHQDKDLGQCCSPGLCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQOKRLTKTSRCGPMWAFNRFVDTWARDDETVLKHILRASMKKLTTRKQGLPPP 360
Db 301 TRIAFKVQOKRLTKTSRCGPMWAFNRFVDTWARDDETVLKHILRASMKKLTTRKQGLPPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPRPIWSASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPRPIWSASPPRARSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLQWNSMPRLPTDLVEGPMFRHYDFRQSCWV 480
Db 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLQWNSMPRLPTDLVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDQALPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLGHL 540
Db 481 RAISOEDQALPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLGHL 540
QY 541 ESSQFPFPGF 549
Db 541 ESSQFPFPGF 549

Handwritten signature: Scott Powers

RESULT 2
US-10-839-882-12
; Sequence 12, Application US/10839882
; Publication No. US20040203106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: CORLEY, Neil C.
; APPLICANT: LAL, Preeti
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: JUNMING, Yang
; APPLICANT: SHIH, Leo L.
; TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
; FILE REFERENCE: PF-0619 PCT
; CURRENT APPLICATION NUMBER: US/10/839,882
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US/09/807,452
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/175,737; unassigned; 60/118,559; 09/249,740; unassigned; 60/154,336
; PRIOR FILING DATE: 1998-10-20, 1998-10-20, 1999-02-04, 1999-04-11; 1999-04-11;
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4325626CD1
US-10-839-882-12

Query Match 99.8%; Score 2978; DB 17; Length 549;

Best Local Similarity 99.8%; Pred. No. 1.6e-224;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVEVAGSMAQEREDITMKYEGHRAGLPEDKPKPFRSYNNVNDHLGIVHETELPPL 60
Db 1 MDVEVAGSMAQEREDITMKYEGHRAGLPEDKPKPFRSYNNVNDHLGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSKWVMDLGDWEKYKSRKLIIDRAYKGMNIRGPMWSVLNTEEM 120
Db 61 TAREAKQIRREISRSKSKWVMDLGDWEKYKSRKLIIDRAYKGMNIRGPMWSVLNTEEM 120
QY 121 KLNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFPRDRYGTQORELHILLAYEY 180
Db 121 KLNPGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFPRDRYGTQORELHILLAYEY 180
QY 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOQE 240
Db 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVOLLASERHSLQGFHSPNGTVOGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGQCCSPGLCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHQDKDLGQCCSPGLCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQOKRLTKTSRCGPMWAFNRFVDTWARDDETVLKHILRASMKKLTTRKQGLPPP 360
Db 301 TRIAFKVQOKRLTKTSRCGPMWAFNRFVDTWARDDETVLKHILRASMKKLTTRKQGLPPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPRPIWSASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPRPIWSASPPRARSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLQWNSMPRLPTDLVEGPMFRHYDFRQSCWV 480
Db 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLQWNSMPRLPTDLVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDQALPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLGHL 540
Db 481 RAISOEDQALPCWOAEHPAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLGHL 540
QY 541 ESSQFPFPGF 549
Db 541 ESSQFPFPGF 549

RESULT 3
US-10-071-838-18
; Sequence 18, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17.codi
US-10-071-838-18

Query Match 99.0%; Score 2952; DB 14; Length 549;
Best Local Similarity 99.1%; Pred. No. 1.7e-222;
Matches 544; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVHETELPPL 60
Db 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVHETELPPL 60
Qy 61 TAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEEM 120
Db 61 TAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEEM 120
Qy 121 KLKNPGRYQIMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKQRELLHLLAYEEY 180
Db 121 KLKNPGRYQIMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKQRELLHLLAYEEY 180
Qy 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
Db 181 NPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
Qy 241 HVVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVVEGEQALMP 300
Db 241 HVVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVVEGEQALMP 300
Qy 301 TRIAFKVOQKRLTKTSRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 360
Db 301 TRIAFKVOQKRLTKTSRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 360
Qy 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPPARFPRPIWSASPPRAPRSSTPCPGGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPPARFPRPIWSASPPRAPRSSTPCPGGA 420
Qy 421 VREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLDVEGWFHRYDFRQSCW 480
Db 421 VREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLDVEGWFHRYDFRQSCW 480
Qy 481 RAISQEDQLAPCWAQAEHPAERVRSFAFAAPSTDSOGTFRARDEQPCAPTSGPCLCGHL 540
Db 481 RAISQEDQLAPCWAQAEHPAERVRSFAFAAPSTDSOGTFRARDEQPCAPTSGPCLCGHL 540
Qy 541 ESSQPPPGF 549
Db 541 ESSQPPPGF 549

RESULT 4
US-10-071-838-4
; Sequence 4, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071, 838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267, 615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
US-10-071-838-4

Query Match 98.4%; Score 2934.5; DB 14; Length 610;
Best Local Similarity 89.8%; Pred. No. 4.6e-221;
Matches 548; Conservative 0; Mismatches 1; Indels 61; Gaps 1;
Qy 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVH----- 53

Db 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVQSCRSWES 60
Qy 54 -----ETELPP 59
Db 61 APQEGCPFPFVPSPLSPLELDRASFPWGSAPRLGPLQAPCCSSALLPGLPYSETELPP 120
Qy 60 LTAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEE 119
Db 121 LTAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEE 180
Qy 120 MKLKNPGRYQIMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKQRELLHLLAYEE 179
Db 181 MKLKNPGRYQIMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQKQRELLHLLAYEE 240
Qy 180 YNPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQ 239
Db 241 YNPEVGYCRDLISHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQ 300
Qy 240 EHVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVVEGEQALMP 299
Db 301 EHVATSQPKTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVVEGEQALMP 360
Qy 300 ITRIAFKVOQKRLTKTSRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 359
Db 361 ITRIAFKVOQKRLTKTSRCGPMWAFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGDLP 420
Qy 360 PAKPEQSSASRPVPASRGKTLCKGDRQAPPPARFPRPIWSASPPRAPRSSTPCPGG 419
Db 421 PAKPEQSSASRPVPASRGKTLCKGDRQAPPPARFPRPIWSASPPRAPRSSTPCPGG 480
Qy 420 AVREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLDVEGWFHRYDFRQSCW 479
Db 481 AVREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLDVEGWFHRYDFRQSCW 540
Qy 480 VRAISQEDQLAPCWAQAEHPAERVRSFAFAAPSTDSOGTFRARDEQPCAPTSGPCLCGHL 539
Db 541 VRAISQEDQLAPCWAQAEHPAERVRSFAFAAPSTDSOGTFRARDEQPCAPTSGPCLCGHL 600
Qy 540 LESSQPPPGF 549
Db 601 LESSQPPPGF 610

RESULT 5
US-10-108-260A-3119
; Sequence 3119, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3119
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3119

Query Match 94.8%; Score 2827; DB 15; Length 527;
Best Local Similarity 95.6%; Pred. No. 1e-212;
Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
Qy 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVHETELPPL 60
Db 1 MDVVEVAGSWWAQEREDIIIMKYEKGRAGLPEDKGPFRSYNNNVNDHLGIVHETELPPL 60
Qy 61 TAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEEM 120
Db 61 TAREAKQIRREISRKSKVMDMLGWDEKYSRKLIIDRAYKGMPMNIRGPMWSVLLNIEEM 120

QY 121 KLKNGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHLLAYEY 180
Db 121 KLKNGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHLLAYEY 180
QY 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGGTVOGLQDOQE 240
Db 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGGTVOGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPKTMGHQ-----ISLGLTLRLMDVYLVEGEQALMPI 278
QY 301 TRIAFKVQQRLLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGLPPP 360
Db 279 TRIAFKVQQRLLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGLPPP 338
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPSSTPCPGA 420
Db 339 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPSSTPCPGA 398
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDLVGEPMFRHYDFRQSCWV 480
Db 399 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDLVGEPMFRHYDFRQSCWV 458
QY 481 RAISOEDQLAPCWOAEHHAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 540
Db 459 RAISOEDQLAPCWOAEHHAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 518
QY 541 ESSQFPPPGF 549
Db 519 ESSQFPPPGF 527

RESULT 6
US-10-094-466-62
; Sequence 62, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spytek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 62
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-094-466-62
Query Match 93.3%; Score 2784; DB 14; Length 549;
Best Local Similarity 93.8%; Pred. No. 2.5e-209;
Matches 515; Conservative 13; Mismatches 21; Indels 0; Gaps 0;
QY 1 MDVVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPFRSYNNVNDHIGVHETELPPL 60
Db 1 MDVVEVAGSWWAQEREDIMKEYEKGRAGLPEDKGPFRSYNNVNDHIGVHETELPPL 60
QY 61 TAREVQIRREISRSKSWVDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
Db 61 TAREVQIRREISRSKSWVKMLGEWDTYKNSRKLIDRAYQGI PMNIRGPMWSVLNIEEI 120
QY 121 KLKNGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHLLAYEY 180
Db 121 KLKNGRYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQKRELLHLLAYEY 180
QY 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGGTVOGLQDOQE 240
Db 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVQLASERHSLQGFHSPNGGTVOGLQDOQE 240
QY 241 HVVATSQPKTMGHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSLPNTMWHQDKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVQQRLLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGLPPP 360
Db 301 TRIAFKVQQRLLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLRASMKKLTTRKQGLPPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPSSTPCPGA 420
Db 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPRPIWSASPPRAPSSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDLVGEPMFRHYDFRQSCWV 480
Db 421 VREDTYPVGTQGVSPALAOGGPQGSWRFLOWNSMPRLPTDLVGEPMFRHYDFRQSCWV 480
QY 481 RAISOEDQLAPCWOAEHHAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 540
Db 481 RAISOEDQPATCWOAEHHAERVRSAFAAPSTDSQGTFRARDEQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPPPGF 549
Db 541 ESSQFPPPGF 549

RESULT 7
US-10-071-838-6
; Sequence 6, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
US-10-071-838-6

```
Query Match      93.0%; Score 2773; DB 14; Length 515;
Best Local Similarity 93.8%; Pred. No. 1.7e-208;
Matches 515; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY      1 MDVEVAGSWAQRERDIIIMKEKGRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
      |||
      1 MDVEVAGSWAQRERDIIIMKEKGRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
      |||
QY      61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
      |||
      61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
      |||
Db      61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
      |||
QY      121 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 180
      |||
      121 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 180
      |||
Db      121 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 180
      |||
QY      181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
      |||
      181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
      |||
Db      181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
      |||
QY      241 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 300
      |||
      241 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 300
      |||
Db      241 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 300
      |||
QY      301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 360
      |||
      301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 360
      |||
Db      267 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 326
      |||
QY      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 420
      |||
      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 420
      |||
Db      327 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 386
      |||
QY      421 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 480
      |||
      421 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 480
      |||
Db      387 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 446
      |||
QY      481 RAISOEDQLAPCWAEPHAEVRVRSAPAASTDSQGTFFRARDQPCAPTSGPCLGHL 540
      |||
      481 RAISOEDQLAPCWAEPHAEVRVRSAPAASTDSQGTFFRARDQPCAPTSGPCLGHL 540
      |||
Db      447 RAISOEDQLAPCWAEPHAEVRVRSAPAASTDSQGTFFRARDQPCAPTSGPCLGHL 506
      |||
QY      541 ESSQFPFPGF 549
      |||
      541 ESSQFPFPGF 549
      |||
Db      507 ESSQFPFPGF 515
      |||

RESULT 8
US-10-071-838-17
; Sequence 17, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TRE-2/USP6
US-10-071-838-17

Query Match      73.5%; Score 2191.5; DB 14; Length 786;
Best Local Similarity 81.6%; Pred. No. 1e-162;
```

```
Matches 408; Conservative 33; Mismatches 58; Indels 1; Gaps 1;

QY      1 MDVEVAGSWAQRERDIIIMKEKGRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
      |||
      1 MDVEVAGSWAQRERDIIIMKEKGRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
      |||
Db      1 MDVEVAGSWAQRERDIIIMKEKGRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
      |||
QY      61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
      |||
      61 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 120
      |||
Db      60 TAREAKQIRREISRSKSKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEM 119
      |||
QY      121 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 180
      |||
      121 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 180
      |||
Db      120 KLKNPGRYQIMKEGKRSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEY 179
      |||
QY      181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
      |||
      181 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 240
      |||
Db      180 NPEVGYCRDLSHIALFLLYLPEEDAFWALVQLLASERHSLQGFHSPNGGTVOGLQDOQE 239
      |||
QY      241 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 300
      |||
      241 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 300
      |||
Db      240 HVVATSQPKTMGHQDKDLGQCSPGCLIRLIDIGISLGLTLRLMDVYLVEGEQALMPI 299
      |||
QY      301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 360
      |||
      301 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 360
      |||
Db      300 TRIAFKVQOKRLTKTSRCGPWFARFCNRFVDTWARDEDTVLKHLRASMKGKLTTRKQGDLPPL 359
      |||
QY      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 420
      |||
      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 420
      |||
Db      360 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPPARFPPIWSASPPRAPRSSTPCPGGA 419
      |||
QY      421 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 480
      |||
      421 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 480
      |||
Db      420 VREDTYPVGTQGVPSPALAQGGPQGSWRFLOWNSMPRLPTDLVEGPFWRHYDFRQSCWV 479
      |||
QY      481 RAISOEDQLAPCWAEPHAE 500
      |||
      481 RAISOEDQLAPCWAEPHAE 500
      |||
Db      480 RAISOEDQLAPCWAEPHAE 499
      |||

RESULT 9
US-10-108-260A-4446
; Sequence 4446, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4446
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4446

Query Match      14.5%; Score 433.5; DB 15; Length 210;
Best Local Similarity 53.3%; Pred. No. 1e-25;
Matches 89; Conservative 26; Mismatches 47; Indels 5; Gaps 2;

QY      18 IMKYEGHRAGLPEDKGP--PPRSYNNVNDHLGIHETELPPLTAREAKQIRREISRK 75
      |||
      18 IMKYEGHRAGLPEDKGP--PPRSYNNVNDHLGIHETELPPLTAREAKQIRREISRK 75
      |||
Db      19 ITTKEGHRAGAALVDLGHQVDVAKXTNN--LGIHETELPPLTAREAKQIRREISRK 75
      |||
QY      76 SKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEMKLKNPGRYQIMKEKG 135
      |||
      76 SKVNDMLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLNIEEMKLKNPGRYQIMKEKG 135
      |||
Db      76 NKWQKMLADWTYKSTKGLSQRVCYIPLAVRGRLSLDLIDIKSQNGKXKXKXKXKXKX 135
      |||
QY      136 KRSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEYNP 182
      |||
      136 KRSEHIQRIIDRVSGTLRKHIIFRDRYGTQKQRELIHLLAYEYNP 182
      |||
Db      136 KRSSRIHICQLDVSHITLQKHMFIQRFQVQKQELCDILVAYSAYNP 182
      |||
```


RESULT 10

```
; US-10-425-115-237632
; Sequence 237632, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 237632
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_148305C.1.pep
US-10-425-115-237632
```

```
Query Match      11.7%; Score 348.5; DB 17; Length 381;
Best Local Similarity 29.6%; Pred. No. 9,8e-19;
Matches 117; Conservative 54; Mismatches 161; Indels 63; Gaps 13;
```

```
OY 33 DKGPFRSYNNVNDHLGI-----VHETELPPLTAREAKQIRREISRK 75
DB 11 EPGPAPSR-----VDRFGFIKPEQGNSPDGI PKGSIHERE-----REERRIR----- 54
OY 76 SKVNDMLG---DWEKY--KSSRKLIDRAYKGMPMNIRGPMWSVLLNIEEMKLNPGRYQ 129
DB 55 -KWRKMIGVGNDWKHYVRNPHVVKRIRKGI PDCLRGLVWQLISGRDLLMNPVYE 113
OY 130 IMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQRELIHLILAYEYNPVEGYCRD 189
DB 114 TLVIYETSASE--LEIRDISRTFSPSHIFQQRHGPQGRSLYNILKAYSVDYDRDVGYYQG 171
OY 190 LSHIALFLVLPBEDAFWALVQLLASERHS-LOGFHSPNGGTVOGLDQOEHVATSQP 248
DB 172 MGFLAGLLLYMSEEDAFWLLVALLKGAVHAPMEGLYQAGLPVQOYLSQFEKLVIELMP 231
OY 249 KTMGHQDKDLCGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPITRIAFKYQ 308
DB 232 KLGQHFVEEMINPSMYASQWFTVFSYSPFHLTLRWMDVFLYEGIKVVFQVGLALLRLC 291
OY 309 QKRLTKTSRCGPWAR--FCNRFVDTWARDEDTVLKLHRSAMKLTTRKQGDLPAPKPEQG 366
DB 292 HDDLVKL-----PFEKLLYALRNFPPEATDPD-VLLPLAFSPKVSRLLELQKEYQKQEG 346
OY 367 ----SSASRPVP-----ASRGKTLCKGDRQ 388
DB 347 TSETSSGERLQPLVSKTMSRVGSRVVSNLTTADRK 381
```

RESULT 11

```
US-10-425-114-64834
; Sequence 64834, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64834
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4720-007-A6_FLI.pep
US-10-425-114-64834
```

```
Query Match      11.7%; Score 348.5; DB 15; Length 434;
Best Local Similarity 30.1%; Pred. No. 1.2e-18;
Matches 116; Conservative 56; Mismatches 157; Indels 57; Gaps 14;
```

```
OY 33 DKGPFRSYNNVNDHLGI-----VHETELPPLTAREAKQIRREISRK 75
DB 64 EPGPAPSR-----VDRFGFIKPEQGNSPDGI PKGSIHERE-----REERRIR----- 107
OY 76 SKVNDMLG---DWEKY--KSSRKLIDRAYKGMPMNIRGPMWSVLLNIEEMKLNPGRYQ 129
DB 108 -KWRKMIGVGNDWKHYVRNPHVVKRIRKGI PDCLRGLVWQLISGRDLLMNPVYE 166
OY 130 IMKEKGRSSEHIQRIDRDVSGTLRKHIFFRDRYGTQRELIHLILAYEYNPVEGYCRD 189
DB 167 TLVIYETSASE--LEIRDISRTFSPSHIFQQRHGPQGRSLYNILKAYSVDYDRDVGYYQG 224
OY 190 LSHIALFLVLPBEDAFWALVQLLASERHS-LOGFHSPNGGTVOGLDQOEHVATSQP 248
DB 225 MGFLAGLLLYMSEEDAFWLLVALLKGAVHAPMEGLYQAGLPVQOYLSQFEKLVIELMP 284
OY 249 KTMGHQDKDLCGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEQALMPITRIAFKYQ 308
DB 285 KLGQHFVEEMINPSMYASQWFTVFSYSPFHLTLRWMDVFLYEGIKVVFQVGLALLRLC 344
OY 309 QKRLTKTSRCGPWAR--FCNRFVDTWARDEDTVLKLHRSAMKLTTRKQGDLPAPKPEQG 366
DB 345 HDDLVKL-----PFEKLLYALRNFPPEATDPD-VLLPLAFSPKVSRLLELQKEYQKQEG 399
OY 367 ----SSASRPVPASRGKTLCK-GDR 387
DB 400 TSETSSGERLQPLA--SKTMSRVGSR 423
```

RESULT 12

```
US-10-425-114-70559
; Sequence 70559, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70559
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE11ID05_FLI.pep
US-10-425-114-70559
```

```
Query Match      11.7%; Score 348.5; DB 15; Length 434;
Best Local Similarity 29.6%; Pred. No. 1.2e-18;
Matches 117; Conservative 54; Mismatches 161; Indels 63; Gaps 13;
```

```
OY 33 DKGPFRSYNNVNDHLGI-----VHETELPPLTAREAKQIRREISRK 75
DB 64 EPGPAPSR-----VDRFGFIKPEQGNSPDGI PKGSIHERE-----REERRIR----- 107
```

```
Db      64 EPGPAPSR-----VDRFGFIKPEQGSPPDGIPKGSIHRE-----REERRIR----- 107
Qy      76 SKVMDLGG----DWEKY--KSSRKLDIDRAYKGMNIRGPMWSVLNIEEMKLNPGRYQ 129
Db      108 -KMRKMIGVGNDWKGIVYRRNPHVKKRIRKGI PDCLRGLVWQLSGSRDLLMNPGYVE 166
Qy      130 IMKEKGRSSEHIQRI DRDVSGLRKHI FPRDRYGTQRELIHLLAYEENPEVGYCRD 189
Db      167 TLVIYETSASE--LEIRDISRTFSPSHIFQQRHGPQGRSLYNILKAYSVDYDRDVGYYQG 224
Qy      190 LSHIALFLIYLPEDAFWALVOLLASERHS-LOGFHSPNGGTVOGLDQDEHVATSO 248
Db      225 MGFLAQLLLLYMSEDAFWLLVALLKGAIVADMEGLYQAGLPLVQGYISQFEKLVIELMP 284
Qy      249 KTMGHQDKDLCGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEALMPITRIAFKVQ 308
Db      285 KLGQHFEEMINPSMYASQWFTVFSYSPFPHLTLRVWDVFLYEGIKVVFQVGLALLRLC 344
Qy      309 QKRLTKTSRCGPWAP--FCNRFVDTWARDEDTVLKLIRASMKKLTREKQGDLPAPPAKPEQG 366
Db      345 HDDLVKL----PFEXLLYALRNFPPEATDPD-VLLPLAFSFKVSSRLLELQKEYQKREQG 399
Qy      367 ----SSASRPVP-----ASRGSKTLCGDRQ 388
Db      400 TSETSSGGRLOPLVSKTMSRVGSRVVSNLTJADRK 434
```

RESULT 13

```
US-10-437-963-139162
; Sequence 139162, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139162
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40481C.1.pep
US-10-437-963-139162
```

```
Query Match      11.0%; Score 329; DB 16; Length 433;
Best Local Similarity 29.8%; Pred. No. 3.9e-17;
Matches 98; Conservative 45; Mismatches 128; Indels 58; Gaps 9;
```

```
Qy      33 DKGPKPRSYNNVNDHLG-----IVHETE-----LPLLTARE 64
Db      11 EPGPAPSRP-----VDRFGFIKPEQGSTPDGISKSKPIHERERESGYLTVHYFTNLTDADI 66
Qy      65 AKQIRREISRSKSWVDLGG----DWEKY--KSSRKLDIDRAYKGMNIRGPMWSVLNIE 118
Db      67 FLGNHRDGRIRKKRKMIGVGSDWKHYVRRNPHVKKRIRKGI PDCLRGLVWQLSGSR 126
Qy      119 EMKLNPGRYQIMKEKGRSSEHIQRI DRDVSGLRKHI FPRDRYGTQRELIHLLAYE 178
Db      127 DLLMNPGVYETLVIYETSASE--LEIRDISRTFSPSHIFQQRHGPQGRSLYNVLKAYS 184
Qy      179 EYNPEVGYCRDISHIALFLIYLPBEDAFWALVOLLASERHS-LOGFHSPNGGTVOGLD 237
Db      185 VYDRDVGYYQGMGFLAGLLLLYMSSEDAFWLLVALLKGAIVADMEGLYQAGLPLVQGYLY 244
```

```
Qy      238 QDEHVATSOPTKMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEQAL 297
Db      245 QFEKLVLEQMPQLGQHFIEMINPSMYASQWFTVFSYSPFPHLTLRVWDVFLYEG----- 300
Qy      298 MPITRIAFKVQQRKLTTSRCGPWAPFCN 326
Db      301 ---IKVFQVGLALL-----RFCH 316
```

RESULT 14

```
US-10-051-874-138
; Sequence 138, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles B
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051, 874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268, 595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325, 306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262, 587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272, 409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262, 454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276, 777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291, 672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330, 336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265, 530
```


OY 283 -----LRLWDVYLVEGEQALMPITRIAFKVCQKRLTKTSRCGPWAFCNRFVDITWA 333
Db 254 ARTLPWASVLRVWDMFCEG-----VKIIFRVALVLRHTLGSVEKLRSCQMYETME 306
OY 334 R-----DEDTVLKHL-----RASMKGKLTROGDL----- 357
Db 307 QLRNLPQOCMOEDFLVHEVTNLPVTEALIERENAAQLKWRBETRGELQYRPSRRLHGSRA 366
OY 358 -----PPPAKEGSSASRPVPASRGKTLCKDRQAPPGPPARFPRIWSASPPRA 409
Db 367 IHEERRRQOPPLGP-SSSLSLPLGLKSRGSR--AAGGAPSP-PPVR-----RA 411
OY 410 PRSSTPCPGAV-REDTYPVGTQGVPSPALAOGGPGS 446
Db 412 --SAGPAPGPVVTAGLHP---SLPSPT-GNSTPLGS 442

Search completed: February 4, 2005, 08:06:45
Job time : 88 secs

This Page Blank (uspio)